



Doctorado en Educación

**Concepciones sobre una evaluación justa
de los estudiantes.**

Una aproximación fenomenográfica

**Fair and just assessment conceptions of
students.**

A phenomenographical approach

Tesis doctoral

Nina Hidalgo Farran

Directores:

F. Javier Murillo Torrecilla

Reyes Hernández Castilla

2 0 1 7

*Als meus pares, per ensenyar-me que l'esforç,
la constància i els somnis van sempre donats de la mà.*

Índice

Agradecimientos	11
Resumen.....	19
Abstract.....	25
Capítulo 1. Introducción	31
Capítulo 2. Organización y desarrollo de la tesis.....	41
2.1. Antecedentes.....	42
2.2. Sentido de la tesis doctoral.....	43
2.3. Presentación de los artículos que componen la tesis	46
2.4. Aportación original en la tesis	54
Capítulo 3. Objetivos de la investigación	57
3.1. Preguntas de investigación.....	58
3.2. Objetivos de investigación.....	59
Capítulo 4. Antecedentes teóricos.....	61
4.1. Justicia Social como prisma para mirar la educación.....	64
4.2. Evaluación de estudiantes.....	74
4.3. Concepciones sobre evaluación de estudiantes	83
4.4. Concepciones sobre una evaluación justa de estudiantes.....	100
Capítulo 5. Metodología.....	117
5.1. Estudio previo: incidencia de la evaluación en el rendimiento.....	118
5.2. Investigación de las concepciones sobre una evaluación justa.....	123
Capítulo 6. Resultados.....	145
6.1. Estudio previo: Incidencia de la evaluación en el rendimiento.....	146
6.2. Investigación de las concepciones. Subestudio 1: Concepciones docentes sobre evaluación	156

6.3. Investigación de las concepciones. Subestudio 2: Concepciones docentes sobre una evaluación justa	166
6.4. Investigación de las concepciones. Subestudio 3: Concepciones de estudiantes sobre una evaluación justa	180
Capítulo 7. Discusión.....	197
7.1. Estudio previo. Incidencia de la práctica evaluativa en el rendimiento de los estudiantes	198
7.2. Las concepciones sobre evaluación.....	201
Capítulo 8. Conclusiones.....	209
8.1. Lecciones aprendidas de la investigación.....	210
8.2. Fortalezas y limitaciones de la investigación.....	213
8.3. Líneas futuras de investigación	216
8.4. Palabras de cierre	220
Capítulo 9. Conclusions.....	223
9.1. Lessons learned of the research.....	224
9.2. Strengths and limitations of the research	227
9.3. Future research lines	230
9.4. Closing remarks.....	233
Referencias.....	237
Anexo 1. Publicaciones que conforman la tesis.....	267
A1.1. Publicación 1. Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España.....	269
A1.2. Publicación 2. Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes.....	295
A1.3. Publicación 3. Incidencia del contexto socio- económico en las concepciones docentes sobre evaluación.....	319
A1.4. Publicación 4. Las concepciones sobre una evaluación justa de los estudiantes. Un estudio fenomenográfico desde la perspectiva de los docentes	353

A1.5. Publicación 5. Students' conceptions about a fair assessment of their learning.....	373
---	-----

Anexo 2. Factor de impacto de las revistas.....383

A2.1. Murillo, F. J., Martínez-Garrido, C. e Hidalgo, N. (2014). Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España. <i>ESE-Estudios sobre Educación</i> , 27, 91-113. doi: 10.15581/004.27.91-113.....	385
A2.2. Hidalgo, N. y Murillo, F. J. (2017). Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes. <i>REICE. Revista Iberoamericana de Calidad, Eficacia y Cambio en Educación</i> , 15(1), 107-128. doi:10.15366/reice2017.15.1.007.....	387
A2.3. Murillo, F. J., Hidalgo, N. y Flores, S. (2016). Incidencia del contexto socio-económico en las concepciones docentes sobre evaluación. <i>Profesorado, Revista de Currículum y Formación del Profesorado</i> , 20(3), 251-281.	389
A2.4. Murillo, F. J. e Hidalgo, N. (en prensa). Las concepciones sobre una evaluación justa de los estudiantes. Un estudio fenomenográfico desde la perspectiva de los docentes. <i>Revista Complutense de Educación</i> , en prensa.....	391
A2.5. Murillo, F. J. e Hidalgo, N. (2017). Students' conceptions about a fair assessment of their learning. <i>Studies in Educational Evaluation</i> , 53, 10-17. doi: 10.1016/j.stueduc.2017.01.001	393

Anexo 3. Otras publicaciones y comunicaciones a congresos395

Índice de tablas

Tabla 4.1. Relación de estudios empíricos sobre una evaluación justa.....	109
Tabla 5.1. Participantes del subestudio 1. Distribución de los docentes en función del contexto, del nivel educativo y de la titularidad del centro.....	131
Tabla 5.2. Características de los docentes participantes en el subestudio 1.....	131
Tabla 5.3. Participantes del subestudio 2. Distribución de los docentes en función del contexto, del nivel educativo y de la titularidad del centro.....	132
Tabla 5.4. Características de los docentes participantes en el subestudio 2.....	133
Tabla 5.5. Participantes del subestudio 3. Distribución de los estudiantes en función del contexto, del nivel educativo y de la titularidad del centro.....	134
Tabla 5.6. Características de los docentes participantes en el subestudio 3.....	135
Tabla 6.1. Resultados de los modelos multinivel contruidos a partir de la introducción de cada variable explicativa en los modelos ajustados.....	148
Tabla 6.2. Resultados de los modelos multinivel finales.....	151
Tabla 6.3. Resumen del proceso de modelado multinivel.....	153
Tabla 6.4. Distribución de los docentes según la frecuencia de realización de diferentes actividades de evaluación. Datos en porcentaje	154
Tabla 6.5. Resumen de las concepciones de los estudiantes acerca de una evaluación justa.....	196

Índice de figuras

Figura 2.1. Sentido de la investigación.....	45
Figura 2.2. Publicaciones que conforman la presente tesis doctoral.....	48
Figura 4.1. Antecedentes teóricos de la investigación	64
Figura 4.2. Enfoque tridimensional de Justicia Social: las tres “R”	67
Figura 5.1. Investigación de las concepciones. Subestudios.....	123
Figura 5.2. Proceso fenomenográfico seguido en el análisis de los datos	139
Figura 6.1. Concepciones de evaluación justa de los docentes.....	180
Figura 7.1. Interrelación entre la actividad de evaluación que desarrolla el docente y la competencia en la que impacta	200

Agradecimientos

Agradecer tantos momentos, tantas palabras, tantas acciones y tanto cariño en estas pocas páginas es complicado, pero quiero aprovechar los próximos párrafos para mencionar a todas las personas que durante estos años han contribuido de una forma u otra en esta tesis doctoral.

Me gustaría empezar dando las gracias al Ministerio de Educación de España por haberme otorgado un contrato de Formación de Profesorado Universitario (FPU). Sin duda, sin este apoyo hubiese sido muy difícil llevar a cabo una investigación de esta envergadura. Gracias por darme la oportunidad de crecer profesionalmente, de formarme y de irme de estancia doctoral a aprender de otras universidades.

De la mano del Ministerio, me gustaría agradecer a toda la Universidad Autónoma de Madrid el haber sido mi casa durante estos años. Parte de mi formación previa también la recibí en esta Universidad, de la que me siento parte y ha sido y

será siempre mi hogar. Gracias a todas las personas con las que me he ido encontrando en este camino: a todos los docentes de la Facultad de Formación de Profesorado y Educación, por todos los aprendizajes que me han regalado; al equipo Decanal, por ser cercano y accesible; al personal de Rectorado, por su inestimable ayuda; al personal de Administración y Servicio de la facultad, por su cariño y amabilidad; y a todos mis alumnos y alumnas con los que he podido aprender a ser mejor docente, compartir experiencias y aprender de todos y cada uno de ellos.

Si esta tesis tengo que agradecerla a alguien es sin duda, y en mayúsculas, a mi director de tesis F. Javier Murillo. Javier, agradecerte todo lo que me has enseñado y aportado es inestimable. Gracias por transmitirme la pasión por la investigación, por dedicarme todo el tiempo que he necesitado para aprender, reflexionar y debatir contigo. Siempre me llevaré conmigo nuestras reflexiones a las ocho de la mañana, nuestros debates y nuestros esquemas en la pizarra. Gracias por no dejarme nunca caer y tener claros mis objetivos como investigadora y docente. Todo este camino juntos ha supuesto un aprendizaje vital que ha marcado y marcará sin duda mi desarrollo profesional. Gracias por la paciencia, los consejos, la entrega y la confianza depositada en mí. Gracias infinitas y eternas.

Esta tesis tampoco habría sido posible sin mi directora, Reyes Hernández-Castilla. Gracias por creer en mí desde mi primer año de psicopedagogía, por darme la oportunidad de formar parte de GICE y por confiar en mí desde siempre. Gracias por creer en mis posibilidades y en hacerme soñar alto. Las clases contigo, las charlas académicas y todos tus consejos han sido vitales para llegar hoy aquí. Tus abrazos, miradas, risas y cariño me han ayudado mucho en este camino. Citándote, que

hayas formado parte de esta tesis ha sido absolutamente “fenomenal”.

Mi más sincero agradecimiento al Grupo de investigación Cambio Educativo para la Justicia Social (GICE). En GICE me he sentido siempre querida, respetada y valorada. Los aprendizajes adquiridos con todas las personas que forman parte del grupo son incalculables y sin duda han contribuido de forma clara a aprender a ser investigadora. Trabajos de campo, seminarios de formación, seminarios de reflexión, reuniones mensuales, celebraciones... tantos y tantos momentos que me han aportado vivencias a nivel profesional y personal que nunca olvidaré. Gracias a Charo, Concha, Paloma, Ángel, Pilar, Gaby, Lourdes, Reyna, Héctor, Alba, Vera, Ariane, Vanessa, Manuel, Natalia, Araceli, Liliana, Antonio, Santiago E., Toni, Santiago... me llevo momentos perennes de cada uno de vosotros.

Pare i mare, ¿com puc agrair tants anys que portem caminant junts? Gràcies per no tallar-me mai les ales, per fer-me sentir la filla més afortunada i feliç de la terra. Les vostres lliçons de vida han guiat tota aquesta tesis, i cada paraula escrita no és sinó un reflex del que vosaltres, tants i tants anys m'heu transmès. Gràcies per l'amor incondicional, pels ànims, pels silencis necessaris, la paciència i els consells. Mai hauria arribat aquí si vosaltres no haguéssiu confiat en mi desde sempre i sé que també serà per sempre. Us estimo.

Mireia, gràcies per no soltar-me la mà ens aquests 26 anys d'amistat. Aquesta tesis reflexa la meva essència la qual és compartida amb tu. Els moments viscuts, la teva ajuda incondicional, la teva veu i la teva mirada m'han acompanyat aquests anys i sé que sempre ho seguiran fent. Edgar, germà de

L' ànima, gràcies per tot l'amor desde fa tants anys. Sempre que t'he necessitat has estat al meu costat i sense les teves rialles, balls, cançons i "t'estimo" no hagués arribat aquí. Gràcies per admirar-me tant i fer-me sentir sempre l'amiga més especial del món. Prometo tornar-te cadascuna de les infinites coses que has fet per mi. Teresa, ¡quina sorpresa ens va donar la vida trobant-nos! Gràcies per ser tan bona persona i tan bona amiga. Tenir-te a la meva vida es un regal i aquestes pàgines están impregnades dels teus ànims, la teva confiança cega en mi, les teves paraules d'amor i les teves rialles incomparables. Més que amics, germans, és una sort tenir-vos a la meva vida, us estimo a tots tres.

Me gustaría agradecer especialmente esta tesis y todos estos años de investigación a mi compañera de trabajo pero también de vida Haylen. Mi querida amiga gracias por todo, por dejarme formar parte de tu vida y por entregarme una amistad absoluta y sincera. Si cierro los ojos y pienso en estos últimos años, sin duda apareces tú en todos los momentos, con tu corazón: noble, valiente y alegre. Gracias por las risas, por los aprendizajes conjuntos, por los llantos compartidos, por tu ayuda inestimable y sobre todo por demostrarme el verdadero significado de la palabra amistad. Incondicional, así eres Haylenis y lo seguirás siendo en mi vida. Gracias!

Quería agradecer de corazón esta tesis a Nieves. Gracias por no soltarme la mano, por confiar en mí desde el primer instante y por iluminar el final del camino de mi tesis. Gracias por tu sonrisa, tu cariño incondicional, tu generosidad, tu bondad y tu mirada llena de admiración. Qué fortuna haberte encontrado y qué felicidad compartir este momento tan importante en mi vida y los que vendrán junto a ti. Gracias por haber elegido escribir la banda sonora de mis días.

Sin duda alguna, no puedo dejar de hablar de mis compañeros de felicidad y de fatiga, de mis intelectuales favoritos, de la gran familia “Gafas” que hemos formado los estudiantes de doctorado de la Facultad de Formación de Profesorado y Educación. El despacho II-102.2 será imborrable en mi mente. Gracias compañeros por tantos momentos: por ayudarnos en la tesis, leernos, formatear documentos, enseñarnos, debatir, compartir comidas, tomar bebidas energéticas y en definitiva, haber hecho más fácil y feliz esta vida doctoral. Separarme de vosotros será complicado, ¡pero lo que la tesis ha unido no lo va a separar una defensa! Gabi, Marian, Ramón, Ani, Sara, Guille, Tania, Lucia Campo, Lucía Hervás, Laura dos, Miguel, Lucía Sánchez, Tote y Vane, sois geniales y siempre os llevaré conmigo.

Dentro de esta gran familia, me gustaría agradecer, con especial cariño, a Irene, Laura, Roci, María y Cynthia. Irene, amiga fiel e incondicional, gracias por agarrarme la mano siempre tan fuerte para no caerme. Tus consejos, ayudas, sonrisas, abrazos y cariño han sido fundamentales en mi tesis. Laura, gracias por ser como eres, por tu espontaneidad y por tu amistad sincera. Conocerte ha sido una suerte inmensa. Gracias por las risas y ánimos siempre en el momento adecuado. Roci, este camino no habría sido lo mismo sin ti. Eres tan humana, generosa, sonriente y bondadosa que solo con tenerte a mi lado me has inspirado durante todo este trabajo. María, gracias por estar siempre a mi lado y confiar en mí, por tus miradas, tus palabras y tu apoyo. Cynthia, gracias por todos los aprendizajes que hemos compartido, por mostrarme el esfuerzo y superación en primera persona y por tus mejores deseos.

El camino recorrido con la tesis no habría sido posible sin Laurita Juliani, Ro, Nacho, Amparito, Marta y Laura. Laurita,

gracias por entregarme tu amistad sincera desde hace tantos años, por ser mi guía emocional en Madrid y por dejarme compartir tantos momentos especiales de nuestras vidas. Ro, compañera y amiga, eres única y siempre has estado a mi lado creyendo en mí y confiando en que éste día iba a llegar. Gracias por tus sonrisas, nuestras conversaciones y ser siempre tan especial conmigo. Nacho, gracias por ser como eres. Tu amistad, tu sonrisa, tu apoyo y tus miradas han acompañado esta tesis. Nuestros bailes, desconexión, risas y momentos compartidos han sido espectaculares. Amparito y Marta, nuestra amistad ha seguido más allá del máster y siempre siento vuestro cariño y sonrisas conmigo. Laura, gracias por estar a mi lado todo este duro camino de tesis, por tu paciencia y tu apoyo. Gracias también a Mari, Lorenzo y Juan por todo vuestro cariño. Rui, muchas gracias por querer siempre lo mejor para mí.

Moltes gràcies a la Núria, l'Óscar, l'Aida, l'Eli i a totes les wapures per estar sempre al meu costat de forma incondicional. La vostra amistad, amor i tendresa m'han guiat aquests anys. Gràcies a la meva família per confiar en mi i estar al meu costat: padrina Paquita, padrí Josep, Jordi, Elena, Dolors, Marc petit, Marc, Joan, Aida, Carlos, Maribel, Jordi, Alba i Maria. Padrina Mercè, sé que desde el cel em dones força i porto el teu somriure grabat al cor. Aquesta tesis és per tu, per ajudar-me a ser qui sóc avui.

No puedo olvidarme en estos agradecimientos de mi estancia doctoral. Thank you so much to John Schostak for all the opportunities, learning and moments gived to me and lived in Metropolitan Manchester University. Your kindness, knowledge and patience were so helpful in my PhD dissertation. David y Edda, muchas gracias por darme la

oportunidad de aprender con vosotros, de asistir a vuestras clases y de ofrecerme apoyo siempre que lo necesité. Laura, Francisco, Francis y Pablo, nunca podré agradecerlos lo suficiente el afecto que me entregasteis. Mi corazón siempre estará en deuda con vosotros por vuestro amor, generosidad y cariño. Gracias también a Nuria y a Raquel por su amabilidad en tierras inglesas.

Gracias a la familia de Zumba por poner las notas de alegría, color y vivacidad a esta tesis. Darwin, tus canciones y tus bailes me han ayudado a levantarme cuando me he caído. Ahinoa, gracias por preocuparte por mí, por tu apoyo y por todo tu cariño. Gracias también a Vanesa por sus palabras de ánimo, y también a Marian, Ana, Carmen, Tavo, Maryi, Txus y Yasmin por preguntar siempre por mi tesis.

Gracias a todas las personas que forman parte de mi vida y a las que quiero con todo mi corazón. Esta tesis, sin vosotros, no tendría sentido.

Resumen

Vivimos en una sociedad profunda y –quizá– crecientemente injusta. El aumento de la brecha entre los que más y los que menos tienen, la terrible crisis de refugiados o la violencia de género son solo la punta del iceberg de los graves casos de opresión, violencia, discriminación, inequidad, xenofobia, machismo y homofobia que se producen cada día en nuestro entorno cercano y en contextos alejados. Ante esta realidad, la educación tiene un papel fundamental que jugar: o perpetúa y legitima las desigualdades manteniendo el *status quo* social existente o, por el contrario, lucha por una sociedad más justa. Pero no hay una respuesta única, cada escuela y cada docente, puede contribuir a una alternativa u a otra. Depende de lo que haga. De esta forma, el currículo, entendido como el conjunto de experiencias que un niño, niña o adolescente vive en su escolarización, juega un papel determinante. Y, dentro de ello, cobra especial relevancia la evaluación.

Efectivamente, la evaluación es un acto marcadamente político. Qué se evalúa, cómo, con quién, para quién y cómo se devuelven los resultados, entre otros, son elementos fundamentales que inciden directamente en su aprendizaje, su desarrollo personal y académico y su vida tanto presente como futura. La forma en que los profesores y profesoras evalúen, a su vez, se ve influida por las concepciones que tengan del proceso evaluativo. Así, las ideas implícitas que los docentes tienen sobre la evaluación determinarán en gran medida su práctica, siendo fundamental conocerlas si queremos mejorar la evaluación en nuestras escuelas y que esta favorezca que luche contra las injusticias sociales. En ese sentido, es fundamental profundizar en la concepción de docentes y estudiantes sobre una evaluación justa.

Numerosos estudios han evidenciado la relación entre las concepciones sobre evaluación y cómo evalúan los docentes, razón por la presencia cada vez mayor de investigaciones sobre esta temática (Álvarez Méndez, 2001; Azis, 2015; Harris y Brown, 2009; Pajares, 1992). Y, más específicamente, en las últimas décadas se están multiplicando los trabajos acerca de las concepciones sobre una evaluación justa de los estudiantes. Los estudios empíricos de Scott, Webber, Lupart, Aitken y Scott (2014), Peeper y Pathak (2008), Pettifor y Saklofske (2012), Andrade y Du (2007), Tierney, Simon y Charland (2011), así como las propuestas teóricas de Suskie (2002) o Picón-Jacome (2013), son un referente al respecto. Los mismos defienden que existen dos formas principales de concebir una evaluación justa: una perspectiva más cercana a la igualdad, defendiendo la necesidad de una práctica evaluativa igual para todos los alumnos y objetiva; y una concepción más próxima a la equidad, defendiendo que la evaluación debe adaptarse a las

características de cada uno de los estudiantes, ser cualitativa y participativa.

El objetivo general de esta investigación es *conocer las concepciones de docentes y estudiantes sobre una evaluación justa*. El mismo se concreta en los siguientes objetivos específicos: a) Determinar la incidencia del contexto socio-económico del centro educativo en las concepciones docentes sobre evaluación de estudiantes; b) Conocer las concepciones de docentes sobre una evaluación justa; y c) Conocer las concepciones de estudiantes sobre una evaluación justa. Para ello se realizaron tres subestudios complementarios, todos ellos utilizando la fenomenografía como metodología.

Además, se realizó un estudio previo cuyo objetivo fue determinar la incidencia de la forma de evaluar en el desempeño académico de los estudiantes de Educación Primaria. Para ello, se realizó una explotación especial de los datos de la Evaluación de Diagnóstico General (EDG) de 2009 mediante un Análisis Multinivel. Así, la muestra está conformada por 28.708 estudiantes y 1.341 docentes en 887 centros escolares españoles. Los resultados reflejan que los estudiantes cuyos docentes realizan en mayor medida controles después de cada tema o lección, corrigen los deberes y cuadernos y valoran el interés y la participación en clase de los estudiantes obtienen mejores resultados. Y, por el contrario, se ha encontrado una relación inversa entre realizar controles escritos trimestrales y el desarrollo de las competencias de los estudiantes.

Para conocer las concepciones de evaluación de docentes y estudiantes hemos utilizado un enfoque fenomenográfico. Este método es el más adecuado para profundizar en el conocimiento de las concepciones (Marton, 1986; Pozo, 2006; Schraw y

Olafson, 2003). Los participantes de este estudio son 73 docentes y estudiantes (10 profesores en el subestudio 1, 31 docentes en el subestudio 2 y 32 estudiantes en el subestudio 3), seleccionados a partir de tres criterios: el nivel educativo del centro (Educación Primaria y Secundaria), su titularidad (pública y privada) y el contexto donde se ubica (favorable o desafiante). El instrumento utilizado para recoger la información fue la entrevista fenomenográfica, con pocas preguntas generadoras. Para analizar los datos, seguimos rigurosamente una aproximación fenomenográfica al discurso de los participantes, siguiendo las seis fases propuestas por Marton (1986).

Los resultados muestran, en primer lugar, que existen importantes diferencias en las concepciones docentes sobre su práctica evaluativa en función del entorno donde trabaje el profesorado. En este sentido, los docentes que trabajan en escuelas situadas en entornos favorables utilizan instrumentos más cuantitativos de evaluación como el examen y priorizan la objetividad y la transparencia. Por el contrario, los profesores de escuelas ubicadas en contextos desafiantes conciben la evaluación como un proceso más holístico, utilizando instrumentos cualitativos de evaluación, teniendo en cuenta aspectos como la actitud o la participación en el aula y considerando las características de cada uno de sus alumnos.

También se ha encontrado que las concepciones sobre una evaluación justa de los docentes están determinadas por el contexto en el que se encuentran las escuelas donde trabajan, siendo los docentes de centros en contextos favorables los que consideran que una evaluación justa es aquella igualitaria y los profesores de escuelas que se encuentran en entornos desafiantes conciben la práctica evaluativa justa como

equitativa. Las características que definen una evaluación justa “igualitaria” son que es cuantitativa, transparente y objetiva, y las ideas implícitas que vinculan la evaluación con la equidad la consideran como aquella que se adecua al contexto de cada estudiante, que mide el progreso y aprendizaje de cada alumno través de pruebas eminentemente cualitativas. Los estudiantes conciben también la evaluación justa como igualitaria o equitativa con las características mencionadas, pero el concebir la evaluación justa desde un polo u otro no viene determinado por el contexto de su escuela.

Los hallazgos encontrados en nuestro estudio están en coherencia con trabajos análogos como los de Scott et al. (2014), Pettifor y Saklofske (2012) o Tierney, Simon y Charland (2011), los cuales también han encontrado concepciones de evaluación justa más cercanas a la igualdad o próximas a la equidad. En coherencia con nuestra investigación, estudios como los de Brown, Hui, Yu y Kennedy (2011) ponen de manifiesto la incidencia del contexto en la forma de concebir la evaluación por parte de los docentes.

El uso de la fenomenografía como forma de aproximarnos a las concepciones y profundizar en la temática de una evaluación justa son dos de las fortalezas de esta tesis. Esta investigación, como cualquier otra, no está exenta de limitaciones. El número de participantes, el carácter eminentemente cualitativo del método empleado o la existencia de concepciones ambiguas, contradictorias o cambiantes de muchos participantes son algunos de los puntos débiles de nuestro estudio. En cuanto a la prospectiva de esta investigación, se podría seguir profundizando en esta temática desde otra perspectiva metodológica. Inspirándonos en la línea de investigación iniciada por Brown y colaboradores (2003, 2008, 2009, 2011)

sería interesante realizar un estudio ex post facto sobre las concepciones de evaluación justa y analizarlas con Modelos de Ecuaciones Estructurales, donde podríamos confirmar, asimismo, la incidencia del contexto en ellas.

Las presiones actuales por la evaluación han hecho que se convierta en uno de los elementos más determinantes del sistema educativo. La realización de evaluaciones externas, cuantitativas e igualitarias fomenta una evaluación clasificadora y jerárquica que solo beneficia a un determinado grupo de estudiantes. Frente a eso, la realidad social, cada vez más desigual, necesita que la evaluación no favorezca esas inequidades sino que, por el contrario, sea consciente de ellas y convierta la evaluación en una práctica centrada en cada uno de sus estudiantes y que favorezca su desarrollo integral. Sólo con una evaluación más consciente, más crítica, más participativa y más equitativa será posible luchar para lograr una educación, y con ello una sociedad, más justa.

Abstract

We live in a deep and –perhaps– an increasingly unfair society. The widening gap between those who have the most and the least opportunities, the terrible refugee crisis or gender violence are only the top of the iceberg of serious cases of oppression, violence, discrimination, inequity, xenophobia, machismo and homophobia, which occur every day in our immediate surroundings and in remote contexts. Faced with this reality, education plays a fundamental role: either perpetuates and legitimizes inequalities maintaining the existing social *status quo* or, on the contrary, strives for a more fair society. However, there is no a single answer, each school and each teacher can contribute to one alternative or to another. It depends on what any person do. In this way, the curriculum, understood as the set of experiences that a child or adolescent lives in their schooling (or school attendance), plays a decisive role. And, the assessment is particularly relevant within this context,

Effectively, assessment is a markedly political act. What is assessed, how, with whom, for whom and how the results are returned, among others, are fundamental elements that directly affect students' learning, their personal and academic development and their life both present and future. The way in which teachers assess is influenced by their assessment conceptions. Thus, the implicit ideas that teachers have about assessment will largely determine their practice, being fundamental to know if we want to improve the assessment process in our schools and the assessment process impact in the fight against the social injustices. In this sense, it is fundamental to deep in the conceptions of teachers and students about a fair assessment.

Numerous studies have demonstrated the relationship between assessment conceptions and how teachers assess their students. For this, the presence of research on this topic has increased (Azis, 2015, Harris and Brown, 2009, Pajares, 1992). Specifically, the researches on the conceptions about fair assessment of students have multiplied in the last decades. Empirical studies such as Scott, Webber, Lupart, Aitken & Scott (2014), Peeper and Pathak (2008), Pettifor and Saklofske (2012), Andrade and Du (2007), Tierney, Simon, & Charland or theoretical proposals of Suskie (2002) or Picón-Jacome (2013), are a reference of this research topic. All these authors argue that there are two main ways of conceiving a fair assessment: a perspective closer to equality, defending the need for an equal assessment practice for all students and objectivity; and a conception closer to the equity, arguing that assessment must be adapted to the characteristics of each student in order to be qualitative and participatory.

The general aim of this study is *to know the conception of teachers and students about a fair assessment*. Its purpose is specified in the following specific goals. With them, we seek to: a) Determine the impact of socio-economic context of the school on the teaching conceptions about students' assessment; b) Know the teachers' conceptions about fair assessment; and c) Know students' conceptions about fair assessment. To achieve these goals, three complementary sub-studies were carried out, all of them using phenomenography as methodology.

In addition, a previous study was carried out to determine the impact of assessment in academic performance of Primary Education students. To achieve this, a special analysis of data of the General Diagnostic Assessment (GDA) of 2009 was performed using a Multi-Level analysis. Thus, the sample was comprised of 28,708 students and 1,341 teachers in 887 Spanish schools. The results reflect that students with teachers who's perform more tests after each topic or lesson, correct homework and notebooks, and value the interest and participation of their students during these classes, achieve better results. Conversely, it was found an inverse relationship between conducting quarterly written tests and the development of student competencies.

To recognize teachers' and students' assessment conceptions we have used a phenomenographical approach. This method is the best-suited method to deepen the conceptions knowledge (Marton, 1986; Pozo, 2006; Schraw, & Olafson, 2003). The participants of this study were 73 teachers and students (10 teachers in sub-study 1, 31 teachers in sub-study 2 and 32 students in sub-study 3), selected from three criteria: school educational level (Primary and Secondary Education), school ownership (public and private) and school context (favourable

or challenging). The instrument used to collect the information was the phenomenographic interview, with a few generating questions. To analyse data, we rigorously followed a phenomenographic approach closer to the discourse of the participants, following the six phases proposed by Marton (1986).

Firstly, results show that there are important differences in the teachers' conceptions about their assessment practice according to the environment where they work. In this sense, those teachers that work in schools with favourable contexts use more quantitative tests such as examination and prioritize objectivity and transparency of assessment. Conversely, those teachers from schools located in challenging contexts consider the assessment as a more holistic process, using qualitative assessment tools and taking into account aspects such as attitude or participation in the classroom of their students and the characteristics of each of them.

In addition, it has been found that teachers' conceptions about a fair assessment are determined by the context of their schools. In that sense, teachers from centres with favourable contexts consider that fair assessment is the egalitarian, while teachers from centres with challenging environments conceive fairness as an equitable assessment practice. The "egalitarian" fair assessment characteristics are quantitative, transparent, and objective. In addition, it is characterized by the implicit ideas that link the assessment with equity, which fits the context of each student and measures the progress and learning of each student through eminently qualitative assessment tools. Students also conceive fair assessment as egalitarian or equitable with the characteristics mentioned above, but these students also consider that fair assessment conceptions from

one pole or another was not determined by the context of their school.

The findings found in our study are consistent with similar studies. For instance, Scott et al. (2014), Pettifor and Saklofske (2012) or Tierney, Simon and Charland (2011), have also found conceptions of fair assessment closer to equality or close to equity. In coherence with our research, Brown, Hui, Yu and Kennedy (2011) highlighted the influence of the context in the way that teachers think about the assessment process.

The strengths of this PhD dissertation are the use of phenomenography as a way of approaching to conceptions and deepening to fair assessment. This research, like any other, has some limitations. The number of participants, the eminently qualitative nature of the method employed or the existence of ambiguous, contradictory or changing conceptions of many participants are some of the weaknesses of our study. As for the prospective of this research, it would be interesting continue in this topic from another methodological perspective. Inspired by the line of research initiated by Brown et al. (2003, 2008, 2009, 2012), it would be interesting to carry out an ex post facto study about fair assessment conceptions and to analyse this with a Structural Equation Models in order to confirm the incidence of the context on the conceptions.

Current pressures for assessment have made it one of the most important elements of educational system. Conducting external, quantitative, and egalitarian assessments encourages a classification and hierarchical assessment that benefits only some determinate group of students. In this sense, social reality, which is increasingly unequal, requires that assessment does not favour these inequities, but rather be aware of them

and make assessment a practice focused on each of its students and its integral development. Only with a more conscious, critical, participatory and equitable assessment will be possible to fight for a just education, and with it, for a more fair society.



CAPÍTULO 1.

Introducción

Las injusticias duelen, y mucho. Pero más doloroso es sentirlas y no hacer nada para cambiarlas. La filósofa política Marion Iris Young (2010), en su obra póstuma *“Responsabilidad por la justicia”*, nos recuerda la corresponsabilidad de todos y todas en el mantenimiento y el desarrollo de una sociedad profundamente injusta, y nos anima a actuar para cambiar.

Desde la educación, la responsabilidad es aún mayor. De un lado, somos conscientes de que el sistema educativo es el mecanismo que utiliza la sociedad por legitimar las desigualdades. Efectivamente, somos nosotros, los y las docentes, quienes, con nuestras decisiones, otorgamos títulos que, a la postre, justificarán diferencias de ingresos. Pero también, muchos de los que nos dedicamos a la educación hemos elegido esta profesión por la convicción de que es el mejor camino para construir una sociedad diferente. Aunque

suene a tópico, las conocidas palabras atribuidas a Nelson Mandela nos acompañan y alientan en nuestro diario quehacer “La educación es el arma más poderosa que puedes usar para cambiar el mundo”.

Nuestra elección personal es cambiar el mundo a través de la educación, y más particularmente a través de la investigación educativa. Una investigación definida por su rigurosidad, por abordar temas transformadores, por combinar la reflexión y la acción, y por intentar que llegue a los y las docentes... verdaderos protagonistas del acto educativo, entre otras características. Estos cuatro elementos han configurado nuestro desarrollo profesional y definen la tesis que aquí se presenta.

El camino elegido ha sido centrar nuestro trabajo en la conformación de una educación que busque, de forma explícita, sistemática y deliberada, una transformación social y el desarrollo teórico-práctico de una Educación para la Justicia Social (Applebaum, 2004, 2005; Cochran-Smith, Shakman, Jong, Terrell, Barnatt y McQuillan, 2009; Connell, 1997; Jornet, Sancho-Álvarez y Bakieva, 2015; Murillo y Hernández-Castilla, 2014; Petrou, Angelides y Leigh, 2009; Sensoy y DiAngelo, 2014; Stevenson, 2007). Con ese compromiso personal y profesional, desde el grupo de investigación de la Universidad Autónoma de Madrid “Cambio Educativo para la Justicia Social” (GICE) llevamos cinco años trabajando y desarrollando investigaciones, intervenciones educativas, acciones de formación y de transferencia. Es en este marco donde nace esta tesis doctoral que ahora presentamos.

La finalidad última de esta investigación es, por tanto, contribuir al desarrollo de una educación orientada a la consecución de una sociedad más justa.

Bajo ese paraguas teleológico, hemos enfocado gran parte de nuestro esfuerzo durante estos años en la evaluación de estudiantes. A pesar de los grandes avances realizados en estos últimos años, especialmente en el ámbito de la teoría, la evaluación de los estudiantes sigue siendo una práctica profundamente injusta. Hoy por hoy, no nos engañemos, contribuye de forma muy eficaz a reproducir y legitimar las diferencias sociales (Janesick, 2008; Jornet, 2016). Como ya apuntamos en un texto previo, las implicaciones de la evaluación son muchas y muy profundas:

Dime cómo evalúas y te diré qué sociedad construyes. La forma que tenemos de evaluar marca inexorablemente a nuestros estudiantes, en la escuela y a lo largo de toda su vida, y con ello se contribuye a crear una sociedad u otra.
(Murillo e Hidalgo, 2015a, p. 5)

Por la necesidad de seguir profundizando en una evaluación que cambie la sociedad, en la presente tesis doctoral nos focalizamos en la evaluación de estudiantes justa. El encuentro fortuito del estudio de Gavin T. L. Brown y Gerrit H. F. Hirschfeld, profesores de la Universidad de Auckland, provocó que nos interesásemos por las concepciones sobre evaluación. En su estudio *“Students’ conceptions of assessment: Links to outcomes”*, Brown y Hirschfeld (2008) determinan que las concepciones influyen directamente en el rendimiento de los estudiantes, por un lado, y en la práctica de evaluación llevada a cabo por los docentes por otro. En este sentido, si conocemos las concepciones tanto de docentes como de estudiantes sobre evaluación podremos mejorar la evaluación desempeñada.

Conjugando la necesidad de una evaluación para cambiar la sociedad y nuestro interés por centrarnos en las concepciones sobre evaluación, la presente investigación tiene como propósito *Conocer las concepciones de docentes y estudiantes sobre una evaluación justa del alumnado.*

Con este marco general se han planteado tres subestudios independientes pero relacionados entre sí, cada uno de los cuales tiene un objetivo específico propio:

- Determinar la incidencia del contexto socio-económico del centro educativo en las concepciones docentes sobre evaluación de estudiantes.
- Conocer las concepciones de docentes sobre evaluación justa.
- Conocer las concepciones de estudiantes sobre evaluación justa.

Para lograrlo, y tras un profundo análisis de los métodos para aproximarse a las concepciones, se ha optado por desarrollar un estudio fenomenográfico que nos permita acercarnos a las ideas implícitas de docentes y estudiantes acerca de qué es una evaluación justa.

Además, se realizó un trabajo previo, que sirvió de punto de partida para esta tesis. El mismo tenía como objetivo:

- Determinar la incidencia de la forma de evaluar en el desempeño académico de los estudiantes de Primaria.

Este estudio previo, metodológicamente consistía en una explotación especial de la base de datos de la Evaluación General Diagnóstica, del Ministerio de Educación de España, abordada mediante Modelos Multinivel.

La presente tesis doctoral se acoge a la normativa de la Universidad Autónoma de Madrid que permite su organización como un compendio de publicaciones. Textualmente:

El doctorando, con la autorización expresa del director/es de la tesis y de la Comisión Académica responsable del programa de doctorado, puede optar por presentar la tesis doctoral como un compendio de publicaciones. (Procedimiento relativo al tribunal, defensa y evaluación de la tesis doctoral en la Universidad Autónoma de Madrid 11/2015, Artículo 8 p. 8)

El primer requisito formal para ello es el número de artículos, su situación editorial y la calidad de las revistas en las que estén o vayan a estar publicados:

Para ello se requiere que tenga publicados o admitidos para su publicación un mínimo de 3 artículos en revistas científicas de reconocido prestigio o en libros editados de importancia justificada. (Procedimiento relativo al tribunal, defensa y evaluación de la tesis doctoral en la Universidad Autónoma de Madrid 11/2015, Artículo 8, p. 8)

En esta tesis presentamos cinco artículos publicados y/o aceptados en revistas con un impacto reconocido en el ámbito de las ciencias sociales y la educación e indexadas en bases de datos y sistemas de evaluación de prestigio (en el Anexo 2 se incluye el detalle completo del impacto de cada una de las revistas donde hemos publicados los artículos que conforman la tesis). Concretamente son:

Murillo, F. J. e Hidalgo, N. (2017). Students' conceptions about a fair assessment of their learning. *Studies in Educational Evaluation*, 53, 10-16.

doi:10.1016/j.stueduc.2017.01.001.

La revista *Studies in Educational Evaluation* está

indexada en Journal Citation Reports (Web of Science) y SCOPUS.

Murillo, F. J., Martínez-Garrido, C. e Hidalgo, N. (2014). Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España. *ESE-Estudios sobre Educación*, 27, 91-113. doi: 10.15581/004.27.91-113.

La publicación *ESE-Estudios sobre Educación* está indexada en SCOPUS.

Murillo, F. J., Hidalgo, N. y Flores, S. (2016). Incidencia del contexto socio-económico en las concepciones docentes sobre evaluación. *Profesorado, Revista de Currículum y Formación del Profesorado*, 20(3), 251-281.

La publicación *Profesorado. Revista de Currículum y Formación del Profesorado* está indexada en SCOPUS y en la base de datos Emerging Sources Citation Index (ESCI) de la Web of Science.

Murillo, F. J. e Hidalgo, N. (en prensa). Las concepciones sobre una evaluación justa de los estudiantes. Un estudio fenomenográfico desde la perspectiva de los docentes. *Revista Complutense de Educación*, en prensa.

La *Revista Complutense de Educación* está indexada en SCOPUS y en la base de datos Emerging Sources Citation Index (ESCI) de la Web of Science.

Hidalgo, N. y Murillo, F. J. (2017). Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes. *REICE. Revista Iberoamericana de Calidad, Eficacia y Cambio en Educación*, 15(1), 107-128.

doi:10.15366/reice2017.15.1.007.

REICE está indexada en la base de datos Emerging Sources Citation Index (ESCI) de la Web of Science.

La normativa de la Universidad aporta, igualmente, algunas indicaciones acerca de la forma en la que debe estar organizada la tesis. Citamos textualmente:

El ejemplar de la tesis ha de cumplir los siguientes requisitos:

a) Incluir una introducción general que presente los trabajos compendiados, que justifique la temática y que explique la aportación original del autor.

b) Incluir un resumen global de los resultados obtenidos, de la discusión de estos resultados y de las conclusiones finales.

c) Se ha de incluir una copia completa de los trabajos publicados o admitidos para su publicación, haciendo constar el nombre de todos los coautores de los trabajos y la referencia completa de la revista en que los trabajos estén publicados o admitidos para su publicación.
(Procedimiento relativo al tribunal, defensa y evaluación de la tesis doctoral en la Universidad Autónoma de Madrid 11/2015, Artículo 8, p. 8)

Con este planteamiento de fondo, hemos dividido este volumen en nueve capítulos antecidos por un **resumen**, en español y en inglés, donde se presentan los principales rasgos de esta tesis doctoral así como la temática que abordamos, la metodología utilizada y los principales resultados y conclusiones.

El primero de los capítulos es la presente **introducción**, donde se justifica la temática y se indican las características formales de esta tesis doctoral. La **organización y desarrollo de la investigación** es el segundo capítulo. Allí se recogen los

antecedentes que fundamentan el estudio que se presenta en este documento. A continuación, se describen los antecedentes que sustentan la tesis así como el sentido de las publicaciones que la conforman y un resumen del contenido de cada una de ellas. En este apartado también se incluye la aportación original de la doctoranda en cada una de las publicaciones.

El tercer capítulo esté destinado a describir los **objetivos** que han guiado la investigación y el problema del que parte nuestro estudio. Los **antecedentes teóricos** del estudio conforma el siguiente capítulo. El mismo se divide en cuatro bloques. En primer lugar se define qué es la Justicia Social como marco de referencia para la educación; en segundo lugar se conceptualiza la evaluación educativa y, en concreto, la evaluación de estudiantes. A continuación se presentan las principales teorías y estudios acerca de las concepciones sobre evaluación y en cuarto y último lugar se analizan los principales estudios sobre qué es una evaluación justa.

El quinto está dedicado a la **metodología** de esta investigación. En el mismo se recoge el enfoque metodológico utilizado, el instrumento de obtención de información, los participantes así como el muestreo utilizado para su selección y por último las categorías y el proceso para el análisis de los datos obtenidos.

Los **resultados** conforma el sexto, donde se presenta de forma conjunta los hallazgos encontrados en nuestro estudio y publicados en los diferentes artículos. En este caso, se describen los resultados siguiendo el orden de los artículos de una forma más completa y detallada. Tras ello, en el séptimo capítulo, se realiza una **discusión** de los hallazgos encontrados con los estudios revisados en el marco teórico.

Las **conclusiones** de la investigación, capítulo centrado mostrar los principales aprendizajes y aportes de esta tesis doctoral para el ámbito educativo, es el octavo. Además se incluye las fortalezas del estudio así como sus limitaciones que presenta. También se reflexiona acerca de las futuras líneas de investigación que pueden derivarse de esta tesis doctoral y se concluye el capítulo con unas palabras a modo de cierre. Como requisito de una tesis con mención internacional, este capítulo se presenta tanto en español como en inglés, que sería el capítulo 9 y último).

Junto con ello se presentan las **referencias** que conforman el presente documento y ultiman esta tesis doctoral, así como los **anexos** que complementan la presente investigación. Allí se incluyen las publicaciones que compendian esta tesis, un análisis del impacto de cada una de las revistas donde han sido aceptados y/o publicados los artículos y una relación de las publicaciones complementarias así como de los congresos a los que hemos asistido durante estos años de formación doctoral.

Investigar acerca de la evaluación de estudiantes es una tarea urgente. La evaluación marca la dinámica del aula, y define, con ello, la forma de enseñar de los docentes y la actitud y papel de los alumnos y alumnas en su aprendizaje. Por ello, es importante conocer qué piensan y que concepciones sobre la práctica evaluativa justa tienen docentes y estudiantes. No hay alternativa: si queremos una educación que favorezca el cambio social es fundamental llevar a cabo una evaluación que trabaje en esa misma dirección.



CAPÍTULO 2.

Organización y desarrollo de la tesis

Como su nombre indica, este segundo capítulo describe la organización de la tesis, aportando el sentido y la contextualización de las publicaciones que la conforman.

De esta forma, en primer lugar, se presentan los antecedentes que han guiado la línea de investigación y que son la base de los artículos compendiados. A continuación, se profundiza en el sentido del estudio que centra la tesis para posteriormente presentar de forma detallada cada uno de los trabajos que la conforman. El último apartado está dedicado a concretar la aportación original de la doctoranda en la tesis en cada uno de los artículos que conforman el trabajo presentado en este documento.

2.1. ANTECEDENTES

La presente tesis doctoral se ha desarrollado entre los años 2013 y 2017 en el marco del Grupo de Investigación *Cambio Educativo para la Justicia Social (GICE)*¹ de la Facultad de Formación de Profesorado y Educación de la Universidad Autónoma de Madrid. El grupo GICE tiene como finalidad última la consecución de una sociedad más justa mediante el desarrollo teórico-práctico de una Educación para la Justicia Social. Para ello, desarrolla investigaciones y realiza acciones de desarrollo, de formación y de transferencia.

La investigación que da lugar a la presente tesis doctoral y el trabajo de continuación que se está realizando se enmarcan en dos proyectos de investigación financiados con fondos del Plan Nacional de Investigación I+D+i que desarrolla el grupo de investigación. El primero de ellos tiene por título “*Escuelas para la Justicia Social*”² y buscaba alcanzar un triple objetivo: conocer las concepciones de Justicia Social de docentes y estudiantes, comprender la cultura de las escuelas que trabajan para la Justicia Social, así como poner en marcha un proceso de transformación en un centro educativo desde planteamientos de Justicia Social. El segundo se titula “*Escuelas en contextos socio-económicamente desafiantes: una aproximación desde la educación para la Justicia Social*”³, y, entre otros objetivos, profundizó en la

¹ <http://www.gice-uam.es/>

² El proyecto “Escuelas para la Justicia Social” (referencia EDU2011-29114) fue financiado por el Ministerio de Ciencia e Innovación a través del Plan Nacional de I+D+i (2012-2014). IP F. Javier Murillo.

³ El proyecto “Escuelas en contextos socio-económicamente desafiantes: una aproximación desde la educación para la Justicia Social” (referencia EDU2014-56118-P) (2015-2017) es financiado con la convocatoria de proyectos de I+D, del programa estatal de fomento de la investigación

influencia del contexto socioeconómico en la cultura de las escuelas.

Dentro de ese marco general el trabajo de la doctoranda se centró en la evaluación de estudiantes como un elemento clave en una enseñanza en y desde la Justicia Social.

De esta forma se continúa y refuerza la línea de trabajo acerca de la Evaluación desarrollada en el marco del área de conocimientos de Métodos de Investigación y Diagnóstico en Educación en el que se encuadran los implicados en esta tesis. Esta línea se materializa no solo en diferentes investigaciones (p. ej., Hernández-Castilla y Salinas, 2008; Murillo, 2006, 2010, 2016; Murillo y Román, 2008, 2009, 2010; Murillo, Román y Hernández-Castilla, 2011; Román y Murillo, 2009) sino también en el trabajo en la edición de la Revista Iberoamericana de Evaluación Educativa.

2.2. SENTIDO DE LA TESIS DOCTORAL

El enfoque de una enseñanza en y desde la Justicia Social junto con el interés por la evaluación de los estudiantes es el marco en el que se desarrolla la presente tesis doctoral.

El primer trabajo que compone la tesis, y que puede considerarse como un estudio preliminar, se centró en determinar el impacto de la evaluación en el desarrollo de los estudiantes (Murillo, Martínez-Garrido e Hidalgo, 2014). Para ello, se realizó una explotación especial de la base de datos de la Evaluación de Diagnóstico General (EDG) de 2009 llevada a cabo por el Ministerio de Educación mediante Modelos

científica y técnica de excelencia por el Ministerio de Economía, Industria y Competitividad. IPs: F. Javier Murillo y Reyes Hernández-Castilla.

Multinivel. Los resultados del mismo fueron claros: la forma en que los docentes evalúan, y especialmente las actividades y estrategias de evaluación que desarrollan, inciden directamente en el aprendizaje de los estudiantes.

Vista esa repercusión, y a partir de un análisis crítico de las prácticas evaluativas de los y las docentes (Murillo e Hidalgo, 2015a, 2015b, 2016a, 2016b), el énfasis se puso en elaborar propuestas alternativas de evaluación más acordes con los planteamientos de Justicia Social que se defienden. De ahí, surgió una propuesta sobre evaluación de estudiantes para la Justicia Social (Hidalgo y Murillo, 2015, 2016; Murillo e Hidalgo, 2015c) en la que en estos momentos están enfocados los esfuerzos. Otro interesante derivación fue la construcción de una escala para medir las actitudes hacia la Justicia Social (Murillo, Hernández-Castilla, Hidalgo y Martínez-Garrido, 2014).

En este camino de construcción, un foco prioritario de atención fue el estudio de las concepciones que tanto estudiantes como docentes tienen tanto de evaluación, en general, como de una evaluación justa. La tesis doctoral, por lo tanto, se centra en los trabajos que han abordado esta temática. En la figura 2.1 se muestra gráficamente ese camino investigador seguido.



FIGURA 2.1. SENTIDO DE LA INVESTIGACIÓN

Fuente: Elaboración propia.

Durante el desarrollo de la investigación, una de nuestras preocupaciones como investigadores era llegar al mayor número de lectores, tanto investigadores como docentes en activo o en formación, posible. En este sentido, las publicaciones realizadas durante estos cuatro años son de diversa índole. Nos hemos preocupado no solo de publicar artículos de alto impacto sino también elaborar trabajos de divulgación en revistas con un gran número de lectores o bien cercanas a los docentes, así como presentar los resultados en diversos eventos o congresos. La relación de las publicaciones y

comunicaciones elaboradas en estos años de desarrollo de la tesis relativa a la evaluación para la Justicia Social se presenta en el Anexo 3.

2.3. PRESENTACIÓN DE LOS ARTÍCULOS QUE COMPONEN LA TESIS

Una vez explicitado el sentido de nuestra investigación pasamos a profundizar en los artículos que se compendian en este trabajo.

Las publicaciones que conforman la tesis doctoral, ya citada en la introducción, son las siguientes:

- Publicación 1. Murillo, Martínez-Garrido e Hidalgo (2014). *Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España.*
- Publicación 2. Hidalgo y Murillo (2017). *Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes.*
- Publicación 3. Murillo, Hidalgo y Flores (2016). *Incidencia del contexto socio-económico en las concepciones docentes sobre evaluación.*
- Publicación 4. Murillo e Hidalgo (en prensa). *Las concepciones sobre una evaluación justa de los estudiantes. Un estudio fenomenográfico desde la perspectiva de los docentes.*
- Publicación 5. Murillo e Hidalgo (2017). *Students' conceptions about a fair assessment of their learning.*

La figura 2.2 presenta el sentido lógico de las mismas.

El primer artículo (P1) es el estudio previo que realizamos donde determinamos la relevancia de la evaluación en el aprendizaje del alumnado. Las otras cuatro publicaciones se centran en el tema de las concepciones sobre evaluación. La publicación 2 es una revisión de investigaciones acerca de las concepciones sobre evaluación; es, por tanto, el marco teórico de esta investigación. La tercera de ellas (P3) recoge un estudio empírico cuyo objetivo es determinar la incidencia del contexto de la escuela donde trabajan los docentes en sus concepciones sobre evaluación. Las dos últimas publicaciones (P4 y P5) profundizan en las concepciones de evaluación justa de docentes y estudiantes respectivamente.

Analicemos uno por uno los cinco artículos que componen esta tesis.

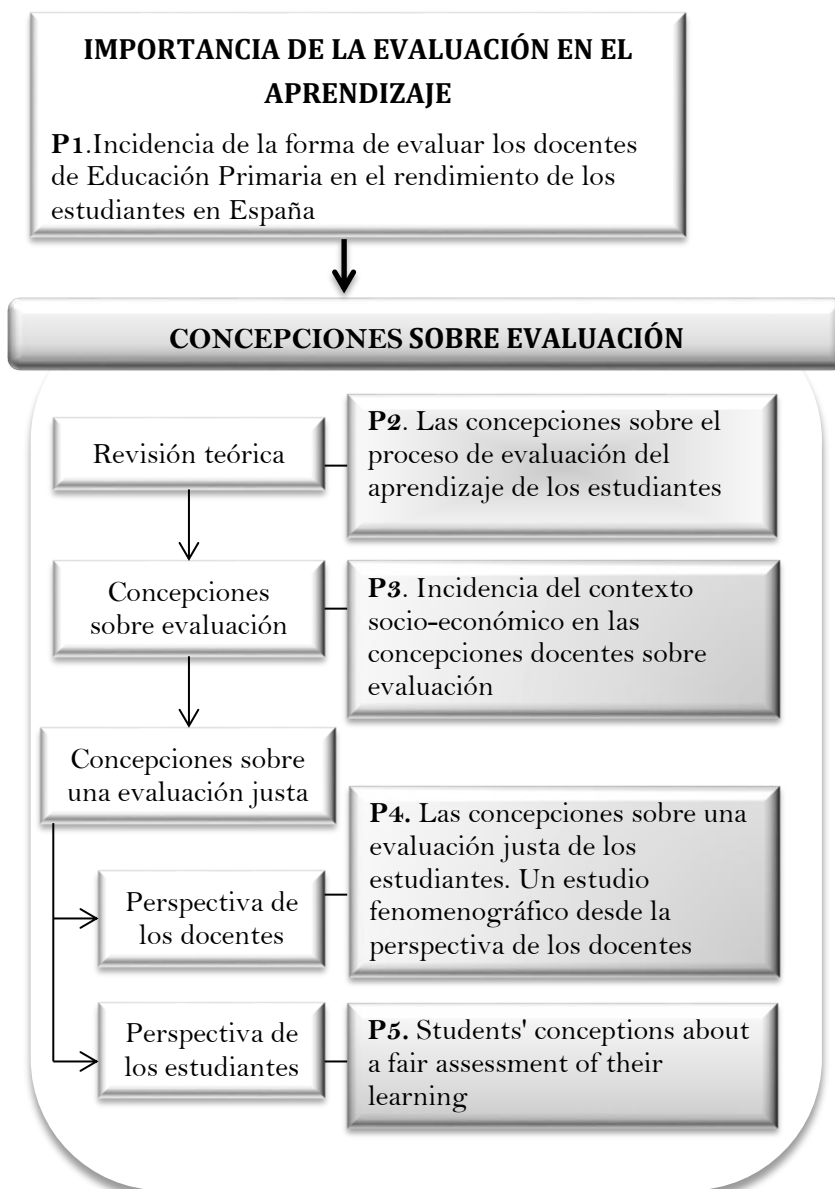


FIGURA 2.2. PUBLICACIONES QUE CONFORMAN LA PRESENTE TESIS DOCTORAL

Fuente: Elaboración propia.

2.3.1. Publicación 1. Murillo, Martínez-Garrido e Hidalgo (2014). Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España

El primer artículo de esta tesis doctoral que, como señalamos, constituye un trabajo previo al estudio de las concepciones, tiene el propósito de determinar la relación entre las actividades de evaluación que llevan a cabo los docentes de Educación Primaria de escuelas españolas y el desarrollo en los alumnos de las Competencias Lingüística, Matemática, Científica y Social y Ciudadana. La idea de la cual parte es clara: la forma en cómo los docentes evalúan a sus estudiantes en el aula incide directamente en su desarrollo cognitivo y socio-afectivo.

Para conocer qué actividades de evaluación incidían más en el desempeño del alumnado, realizamos una explotación especial de los datos de la Evaluación de Diagnóstico General (EDG) de 2009 de España mediante el uso de Modelos Multinivel de tres de niveles (estudiante, aula y escuela). La muestra del estudio es 28.708 estudiantes y 1.341 docentes en 887 centros educativos españoles.

Los resultados muestran, en primer lugar, que la forma de evaluar por parte de los y las docentes de Educación Primaria en España está directamente relacionada con el desempeño académico de los estudiantes. En segundo lugar, analizando qué prácticas de evaluación inciden más en su aprendizaje, los docentes que realizan en mayor medida controles después de cada tema o lección, corrigen los deberes y cuadernos y valoran el interés y la participación en clase de los estudiantes promueven unos mejores resultados por parte de los estudiantes. Y, por el contrario, se ha encontrado una relación

inversa entre realizar controles escritos trimestrales y el desarrollo de las competencias de los estudiantes.

El artículo completo se encuentra en el Anexo 1.1 de la presente tesis doctoral.

2.3.2. Publicación 2. Hidalgo y Murillo (2017). Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes

El segundo artículo que conforma la tesis es el marco teórico de la misma. Su objetivo es, por tanto, profundizar en los estudios actuales centrados en conocer las concepciones sobre evaluación.

La evaluación de los estudiantes es una de las prácticas docentes que más inciden en el aprendizaje y desarrollo de los estudiantes, especialmente porque determina su vida presente y futura. Así, las decisiones que se toman a partir de la evaluación clasifican a los estudiantes y esto influye en su vida académica, profesional y social. Por todo ello, es fundamental conocer las concepciones de evaluación, razón por la cual esta línea de investigación va adquiriendo fuerza en las últimas décadas.

De esta forma, el artículo está estructurado en tres partes. En primer lugar, se ahonda en la definición de las concepciones y su concreción en el ámbito educativo y evaluativo. En segundo lugar, se revisan las investigaciones de concepciones sobre evaluación en cuatro direcciones:

- Los tipos de concepciones sobre evaluación así como las diferencias de dichas ideas implícitas en distintos países que se realizó el mismo estudio.

- Los factores que inciden en la diversidad de concepciones sobre evaluación en docentes y estudiantes.
- La relación entre las concepciones sobre evaluación y la práctica evaluativa llevada a cabo por los profesores y profesoras.
- Otros estudios relacionados con las concepciones sobre evaluación.

Se cierra el artículo con unas reflexiones acerca de lo que hemos aprendido de los estudios analizados relacionados con las concepciones sobre evaluación así como la propuesta de futuras líneas de investigación centradas en aquello que aún nos queda por saber de las creencias implícitas sobre la evaluación del aprendizaje de los estudiantes.

En el Anexo 1.2 se incluye esta publicación.

2.3.3. Publicación 3. Murillo, Hidalgo y Flores (2016). Incidencia del contexto socio-económico en las concepciones docentes sobre evaluación

La tercera publicación que compendia esta tesis recoge un estudio empírico cuyo objetivo es determinar si las concepciones de evaluación del profesorado se ven determinadas por el contexto en el que se encuentra la escuela.

Para ello se ha desarrollado un estudio cualitativo de carácter fenomenográfico. Los participantes son diez docentes de Educación Primaria y Secundaria, la mitad trabajando en un contexto socio-económico favorable y la otra mitad en un contexto desafiante. El instrumento utilizado para la recolección de datos fue la entrevista fenomenográfica.

Los resultados muestran importantes diferencias en las concepciones sobre los referentes de evaluación, sobre los instrumentos usados y su adaptación, y sobre la devolución de la información. Lo más relevante es la diferente concepción que tienen de aquello que es evaluación justa. Para los docentes de entornos favorables una evaluación justa está ligada a la objetividad y a la transparencia, los instrumentos y los criterios de corrección; mientras que para los docentes que trabajan en contextos desafiantes está más relacionada con la adaptación y la medición integral de los avances del estudiante. La conclusión del estudio es la importancia de tener en cuenta el contexto socio-económico, tanto para desarrollar investigaciones como para poner en marcha políticas de apoyo a los centros.

El texto de este estudio se encuentra en el Anexo 1.3 de este volumen.

2.3.4. Publicación 4. Murillo e Hidalgo (en prensa). Las concepciones sobre una evaluación justa de los estudiantes. Un estudio fenomenográfico desde la perspectiva de los docentes

La cuarta publicación presenta una investigación que se centra en conocer las concepciones de los docentes acerca de qué es una evaluación justa. En la revisión teórica del artículo analizamos las concepciones de los profesores, entendidas como las creencias implícitas que tienen sobre el proceso de enseñanza-aprendizaje. Estas concepciones tienen una fuerte incidencia sobre sus prácticas evaluativas y sobre el desarrollo de los estudiantes. Para conocer en profundidad las ideas implícitas de evaluación de los profesores, analizamos los principales estudios de concepciones sobre evaluación desde la

perspectiva de los docentes, tanto internacionales como nacionales.

De las aportaciones de los estudios analizados nace el planteamiento de nuestro artículo. Para lograr nuestro objetivo, hemos desarrollado un estudio cualitativo de carácter fenomenográfico. Los participantes son 31 docentes de Educación Primaria y Secundaria de la Comunidad de Madrid. Los resultados muestran que las concepciones de evaluación justa de los docentes están divididas, estando algunas más cercanas al principio de igualdad y otras al de equidad. La conclusión final del estudio es la influencia del contexto del centro educativo en la concepción de evaluación justa del docente y la relevancia de conocer las concepciones de los profesores para poder implementar prácticas evaluativas que logren una educación y una sociedad más justa.

El artículo se recoge en el Anexo 1.4.

2.3.5. Publicación 5. Murillo e Hidalgo (2017). Students' conceptions about a fair assessment of their learning

El último de los artículos que conforman esta es una investigación que busca conocer las concepciones sobre una evaluación justa de los estudiantes de Educación Primaria y Secundaria. Como puede comprobarse, esta investigación es análoga a la presentada en la cuarta publicación pero centrándonos en este caso en la perspectiva del alumnado.

Para alcanzar ese objetivo se ha llevado a cabo una investigación fenomenográfica con 32 estudiantes. Los resultados muestran que, en términos generales, existen dos tipos de concepciones. Por un lado, una concepción de evaluación justa relacionada con igualdad, transparencia, objetividad, y evaluación de los contenidos; es decir una

concepción igualitaria de la evaluación justa. Por otro, un concepto asociado a ideas tales como adaptación y diversificación de las pruebas, evaluación cualitativa, del esfuerzo y de actitudes... que definen un concepto de evaluación justa ligada a la equidad. Ambas concepciones están asociadas a una concepción de justicia más cercana a la justicia “legal” o a la justicia “social”. Es inevitable relacionar estas concepciones de evaluación justa con el devenir evaluativo del sistema evaluativo español.

El artículo presentado se encuentra en el Anexo 1.5.

2.4. APORTACIÓN ORIGINAL EN LA TESIS

Como no puede ser de otra forma, la presenta tesis doctoral ha sido desarrollada esencialmente por la candidata, Nina Hidalgo Farran, apoyada por sus directores de tesis. Más concretamente los cinco artículos que la componen han sido elaborados y firmados por la candidata junto al director de tesis F. Javier Murillo. En dos de las publicaciones han colaborado, además, dos compañeras doctorandas del grupo de investigación GICE: Cynthia Martínez-Garrido (P1) y Sofía Flores (P3).

En la primera de las publicaciones la aportación de la doctoranda se centró en la revisión de la literatura, redacción de resultados, la elaboración de las conclusiones y la revisión del texto.

En el segundo artículo la doctoranda realizó la búsqueda de las investigaciones más relevantes en el estudio de las concepciones sobre evaluación. Una vez seleccionados los textos, se analizaron en profundidad detectando distintos tipos de estudios dentro de la línea de investigación. Posteriormente,

la doctoranda, en colaboración con el director de la tesis, redactó el artículo teórico a partir de los distintos apartados que lo componen. Por último, se eligió la revista a la que enviar el escrito y se adecuó en cuanto a normas de publicación y estilo.

Las aportaciones de la doctoranda en la tercera, cuarta y quinta publicación son las siguientes:

- *Profundización en la literatura* más relevante acerca de la fenomenografía para conocer mejor cómo diseñar un estudio de ese tipo y qué instrumentos utilizar.
- *Diseño del estudio fenomenográfico*, especialmente en la elaboración y validación del instrumento y la selección de los informantes clave.
- *Contacto* con los centros educativos y realización del trabajo de campo del estudio, compuesto por un total de 73 entrevistas fenomenográficas.
- *Transcripción completa* de las entrevistas y análisis de las mismas con el software Atlasti 7.
- *Redacción de los artículos*. Redactar, en colaboración con el profesor F. Javier Murillo, cada uno de los apartados que componen los tres artículos (introducción, revisión de la literatura, métodos, resultados y discusión) de la presente tesis doctoral.
- *Elección y envío de los artículos a las revistas*. La doctoranda se encargó de elegir y enviar el artículo a los editores de las revistas, proponer evaluadores cuando la revista lo solicitaba así como de responder a los comentarios de cada uno de los revisores y editores.

De esta forma, el trabajo de la doctoranda en la investigación ha sido imprescindible y fundamental para el diseño de las investigaciones, desarrollo del trabajo de campo y redacción de los artículos que componen la presente tesis doctoral.



CAPÍTULO 3.

Objetivos de la investigación

En este capítulo mostramos las preguntas que han guiado esta investigación, ofreciendo los interrogantes claves a los que queremos dar respuesta en este estudio.

Asimismo, presentamos el objetivo general de esta investigación así como los objetivos específicos que han ayudado a concretar los distintos artículos que conforman esta tesis doctoral.

3.1. PREGUNTAS DE INVESTIGACIÓN

El problema de investigación generador de esta investigación es *¿qué concepciones tienen docentes y estudiantes sobre evaluación justa?*

Para dar respuesta al problema planteado, nos formulamos las siguientes preguntas más concretas:

- ¿Cómo conciben los docentes la evaluación de los estudiantes?
- ¿Incide el contexto en el que trabajan los docentes en esas concepciones de cómo evalúan?
- ¿Qué es para los profesores y profesoras una evaluación de estudiantes justa?
- ¿Cómo conciben los estudiantes una práctica evaluativa justa?

Además, como se ha señalado, se desarrolló un estudio previo cuya pregunta generadora fue:

- ¿La forma de evaluar por parte de los docentes de Educación Primaria en España influye en el desempeño académico de sus estudiantes?

A partir de esta serie de preguntas de investigación se formularon los objetivos del estudio, tanto el general como los específicos.

3.2. OBJETIVOS DE INVESTIGACIÓN

El objetivo general que ha guiado esta investigación es:

Conocer las concepciones de docentes y estudiantes sobre una evaluación justa de los estudiantes.

Para lograr dicho objetivo general nos planteamos los siguientes objetivos específicos:

- Determinar si el contexto socio-económico del centro educativo en el que trabajan determina diferencias en las concepciones.
- Conocer las concepciones de Evaluación Justa de docentes de Educación Primaria y Educación Secundaria.
- Conocer las concepciones de Evaluación Justa de estudiantes de Educación Primaria y Educación Secundaria.

A ellos se añade uno previo:

- Determinar la relación entre la forma de evaluar por parte de los docentes de Educación Primaria en España y el desempeño académico de sus estudiantes.



CAPÍTULO 4.

Antecedentes teóricos

La evaluación de estudiantes es una de las actividades educativas más controvertidas y polémicas. El debate que subyace a la práctica evaluativa nace, por un lado, de la complejidad de medir adecuadamente el aprendizaje de cada estudiante (Rodríguez-Gómez, Ibarra-Sáiz, Gallego-Noche, Gómez-Ruiz y Quesada-Serra, 2012; Zabalza, 2003) y, por otro, de las fuertes repercusiones sociales y políticas que tienen las decisiones tomadas a partir de dicha evaluación (Boud y Falchikov, 2007; Delandshere, 2010; Greene, 2006; Lacueva, 1997; Moreno Olivos, 2011, 2014; Murillo e Hidalgo, 2014, 2015a; Pérez Expósito y González Aguilar, 2011).

Sambel, McDowell y Brown (1997), evidenciando la mala reputación de la práctica evaluativa, concluyeron que la

evaluación del aprendizaje de los estudiantes es percibida de forma negativa por los distintos agentes del sistema educativo. Así, en su estudio, encontraron que profesorado y alumnado concebían la evaluación como una molestia, un mal o una injusticia necesaria con consecuencias sociales inequitativas. Quizás por su mala reputación no es extraño que se estén multiplicando en los últimos años las investigaciones que buscan profundizar en distintos aspectos de la práctica evaluativa. Entre las múltiples direcciones que están tomando los estudios sobre evaluación, una de las más fructíferas es aquella que busca conocer las concepciones de docentes y estudiantes sobre evaluación (Atjonen, 2014; Azis, 2012; Brown, 2003, 2004; Brown y Hirschfeld, 2008; Brown y Remesal, 2012; Brown, Lake y Matters, 2011; Chen, Morris y Mansour, 2015; Fives y Gill, 2014; Harris y Brown, 2008, 2009; Rodríguez-Gómez, Ibarra-Sáiz e Ibáñez, en prensa; Winterbottom, Bekker, Conner y Mooney, 2008).

La idea fundamental de la que parte la línea de investigación que busca conocer las concepciones sobre evaluación de los protagonistas del proceso educativo, es que si profundizamos en las concepciones de los docentes podremos entender mejor su práctica evaluativa y mejorarla, y con ello, el desarrollo de los estudiantes (Griffiths, Gore y Ladwig, 2006; Harris y Brown, 2009; Pope, Green, Johnson y Mitchell, 2009; Remesal, 2006, 2011). Asimismo, conociendo las concepciones sobre evaluación de los estudiantes comprenderemos mejor sus vivencias y experiencias con la evaluación y cómo cambiarla para que ésta sea significativa para el alumnado (Brown, 2008; Murillo e Hidalgo, 2015c; Peterson y Irving, 2008; Tierney, 2013; Tierney, Simon y Charland, 2011).

Dentro de la línea de investigación de concepciones sobre evaluación cobran especial relevancia aquellos estudios que se centran en determinar las ideas implícitas acerca de que es una evaluación justa. Definir qué es una evaluación justa no es una tarea fácil, ya que evaluar tiene implicaciones éticas que influyen enormemente en el aprendizaje de los estudiantes (Aitken, 2011; Cline, 2000; Darling-Hammond, 1994; Gilliland, 1993; Lam, 1995; Picón-Jácome, 2013; Pettifor y Saklofske, 2012; Rogers, 2010; Suskie, 2002; Scott et al., 2014). Pese a su importancia, las investigaciones empíricas sobre las concepciones de los docentes sobre una evaluación justa son muy escasas, siendo necesario profundizar en ellas.

Este apartado describe de forma resumida los principales antecedentes teóricos que sustentan la presente tesis doctoral. Del mismo modo que hemos dado sentido a los artículos que compilan esta tesis, los antecedentes teóricos se vertebran en cuatro ejes. En este sentido, y sin el propósito de ser exhaustivos, pasamos a plantear los principales aspectos teóricos trabajados en los distintos estudios realizados que pueden encontrarse en la revisión de la literatura de cada una de las publicaciones.

En primer lugar esbozamos brevemente que entendemos por Justicia Social y cómo se entiende desde una perspectiva educativa. En segundo lugar, consideramos fundamental abordar qué se entiende por evaluación, especialmente la evaluación de estudiantes. Posteriormente, sabiendo que las concepciones influyen directamente en la práctica, es fundamental conocer qué son y cuáles son las principales investigaciones que han profundizado en las ideas implícitas de evaluación de docentes y estudiante. Por último, una de las líneas de investigación más fructíferas dentro de las

concepciones sobre evaluación es aquella que profundiza en cómo es una evaluación justa, exponiendo las principales perspectivas teóricas y los estudios empíricos más relevantes (figura 4.1).



FIGURA 4.1. ANTECEDENTES TEÓRICOS DE LA INVESTIGACIÓN

Fuente: Elaboración propia.

4.1. JUSTICIA SOCIAL COMO PRISMA PARA MIRAR LA EDUCACIÓN

En este primer apartado buscamos, por un lado, definir el concepto de Justicia Social, delimitando en primer lugar qué no es y a qué nos referimos al utilizarlo. Por otro lado, trasladamos este concepto al ámbito educativo analizando qué es una educación socialmente justa.

4.1.1. ¿Qué no es Justicia Social?

Antes de definir qué es Justicia Social consideramos importante delimitar qué no es. Habitualmente el término se asocia a otros conceptos que aun estando relacionados son mucho más restrictivos y no muestran la complejidad y matices de lo que es, desde nuestra perspectiva, la Justicia Social.

Cuando hablamos de Justicia Social no nos referimos a (Murillo y Hernández-Castilla, 2014):

- *Derechos Humanos.* Aun siendo un aspecto fundamental que supone que todas las personas tengan un mínimo de dignidad y calidad de vida, no es para nosotros un fin en sí mismo, sino un punto de partida para lograr una sociedad más justa. Desde los organismos públicos es necesario que velen para que todas las personas cuenten con los derechos mínimos irrenunciables.
- *Igualdad de Oportunidades.* Tal como afirma Dubet (2011) la igualdad de oportunidades es un mecanismo perverso que no logra la transformación social sino que sigue manteniendo la diferencia de clases sociales y no permite la movilidad social entre las distintas clases. Así, no hay nada más injusto que todas las personas puedan ser igual de pobres o ricas dándoles a todas lo mismo, cuando las necesidades de las personas son muy diferentes en función de su situación social. La igualdad de oportunidades es, por lo tanto, una falacia que no logra mejorar la situación social de las personas en mayor desventaja y que por lo tanto no se ajusta a la finalidad de la Justicia Social.
- *Distribución equitativa de bienes.* Aun siendo un aspecto muy importante de la Justicia Social, dar más a aquellos

que más lo necesitan no resuelve todas las injusticias que ocurren en nuestro entorno. Por desgracia, muchas de las discriminaciones existentes no son únicamente por temas o aspectos económicos sino también por temas de cultura, capacidad, lengua, religión, etnia, raza, u orientación sexual. Por ello, es necesario complementar la distribución de los bienes entre las personas con el reconocimiento de cada uno de los individuos y la participación en la sociedad como principios añadidos.

- *Desarrollada únicamente dentro del Estado-Nación.* Entender la Justicia Social dentro de cada Estado o delimitación territorial es profundamente injusto, especialmente en los tiempos globalizadores en los que nos encontramos. La Justicia Social debe traspasar las fronteras y lograr una justicia global en todo el mundo (Fraser, 2008).

4.1.2. Justicia Social como concepto dinámico

Una vez clarificado que no entendemos por el término, en la presente tesis se adopta un enfoque tridimensional de la Justicia Social. El concepto Justicia Social no es un concepto estático, consensuado y cerrado; quizá todo lo contrario, es un concepto dinámico, que depende y se afianza en relación a la sociedad en la que se enmarca y de cómo comprenda esta la educación y su papel en la construcción de una ciudadanía democrática (Murillo y Hernández-Castilla, 2011, 2014).

En coherencia con Nancy Fraser (2008), entendemos Justicia Social a partir de tres dimensiones complementarias: la Redistribución, el Reconocimiento y la Participación, para

comprender de forma holística y completa la Justicia Social (figura 4.2).

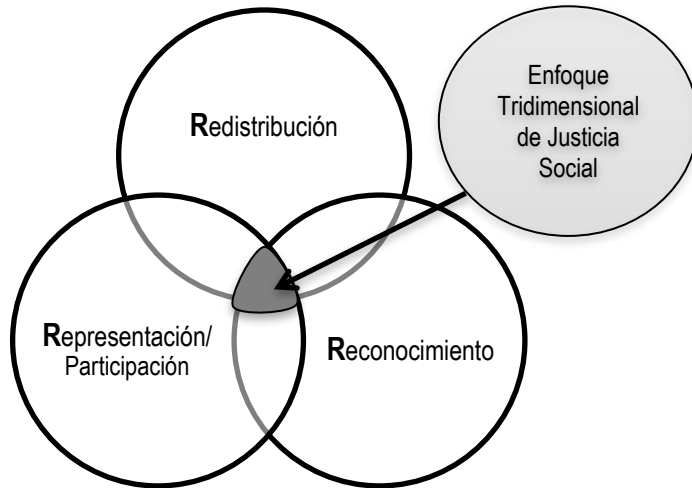


FIGURA 4.2. ENFOQUE TRIDIMENSIONAL DE JUSTICIA SOCIAL: LAS TRES “R”
Fuente: Elaboración propia.

Estas tres dimensiones no son independientes sino que están estrechamente relacionadas y se conjugan para comprender de forma compleja qué es la Justicia Social:

- *Redistribución o Justicia Social redistributiva.* Este principio se basa fundamentalmente en los planteamientos de John Rawls (1971, 2002) el cual defiende la Justicia Social como una distribución equitativa de los bienes primarios (como se ha expuesto en este capítulo). Desde el principio de Redistribución consideramos que las instituciones (entre ellas la educación) deben luchar por la Justicia Social dando más a aquellos que más lo necesitan, compensando así los “azares de la naturaleza” y con ello las desigualdades

creadas por esas diferencias. Asimismo, estamos de acuerdo con Rawls (1971) al afirmar que estas instituciones si no son justas no suponen ningún bien para la sociedad y por lo tanto tienen que ser abolidas.

- *Reconocimiento o Justicia Social centrado en el reconocimiento social y cultural.* Basándonos en el enfoque de Fraser (2008), Honneth (2003, 2007), Taylor (2003) y Young (2000) consideramos que reconocer explícitamente las diferencias de todos los ciudadanos y ciudadanas supone una ausencia de dominación cultural, no reconocimiento o actitud irrespetuosa ante cualquier persona. A partir de estos principios consideramos fundamental valorar social y culturalmente a todos los sujetos así como las diferentes formas de ser, pensar, entender y hacer de todas las personas.
- *Participación o Justicia Social representativa.* Por último, para poder desarrollar una Justicia Social adecuada es fundamental que las instituciones creen las condiciones necesarias para que los ciudadanos y ciudadanas participen plenamente en todos los estamentos de la vida social, especialmente de aquellos sujetos que han sido tradicionalmente oprimidos (Fraser, 2008 y Young, 2000).

Complementando la perspectiva tridimensional, es fundamental lograr una Justicia Estructural, eliminando cualquier tipo de opresión que permita la correcta participación de todos los individuos. Estas cinco caras de la opresión que hay que eliminar de las estructuras sociales y de la sociedad son la explotación, marginación, carencia de poder, imperialismo cultural y violencia (Young, 2000).

Las instituciones existentes en la sociedad tienen que tener la capacidad de ser críticas con la situación de injusticia latente y ser críticas con ella. Pero no tan solo las instituciones, sino también todos los ciudadanos y ciudadanas, quienes tenemos una responsabilidad personal por el logro de una mayor justicia en nuestra sociedad (Young, 2010). Las instituciones y los ciudadanos tienen que ser conscientes de que trabajar por la Justicia Social empieza por las decisiones que se toman, la denuncia de las situaciones injustas y la lucha activa contra dichas injusticias estructurales. Una de las instituciones sociales que más aporta en la identificación de las injusticias y lucha activa contra ellas es la educación.

4.1.3. El papel de la educación en la consecución de una sociedad más justa

Frente a las injusticias que vivimos actualmente en la sociedad, la educación es uno de los procesos más relevantes para compensar esas injusticias y promover la movilidad social. En palabras de Murillo y Hernández-Castilla (2011):

Es necesario recuperar el concepto de Justicia Social en la educación, últimamente sustituido por equidad educativa - mucho más restrictivo y menos ambicioso-, y situarlo en primer término de la agenda política y social. Ya no es suficiente que los niños, niñas y adolescentes aprendan Lengua o Matemáticas, como nos recuerda PISA, es preciso que las escuelas contribuyan a la construcción de una sociedad más justa. (p. 5)

Las escuelas, como instituciones sociales que pueden promover el cambio social, son ámbitos profundamente ideológicos y políticos (Giroux, 1990; Hackman, 2005; Kohl, 2000; McLaren, 1984, 1995, 2010; Sancho-Álvarez, Jornet y González-Such, 2016; Sancho-Álvarez, Jornet y Perales Montolío, 2014). En

palabras de Giroux (2001), este carácter puramente político de la educación se caracteriza por:

Enseñar a los alumnos a afrontar riesgos, oponerse a quienes tienen el poder, ensalzar las tradiciones críticas y reflexionar sobre el modo de utilizarse la autoridad en el aula y en otros lugares pedagógicos. La educación política no sólo da a los alumnos la oportunidad de expresarse críticamente, sino también de alterar la estructura de participación y el horizonte del debate mediante el que se construyen sus identidades, valores y deseos. La educación política establece unas condiciones pedagógicas que permitan comprender a los alumnos la forma en que el poder actúa sobre ellos, mediante ellos y para ellos al servicio de la construcción y expansión de sus funciones como ciudadanos críticos. (p. 26)

Teniendo en cuenta el carácter político de la educación, Giroux (1990, 1992, 2001) considera que la educación persigue tres propósitos fundamentales:

- Analizar la cultura, la ideología y los valores relacionados con el rol del docente.
- Favorecer una relación más estrecha de la educación con la sociedad.
- Desarrollar una política cultural.

Como sucede en la sociedad, en las escuelas, los grupos dominantes y subordinados se entremezclan, donde luchan por tener su espacio. Las luchas de poder que existen en la escuela pueden ser legitimadas por la escuela o bien denunciadas. Las escuelas pueden ser promotoras de las injusticias sociales pero también pueden ser instituciones de denuncia y resistencia. Para que la escuela luche para lograr una mayor justicia Social tiene que impactar en la sociedad: no solo denunciar las situaciones sino cambiarla (Beane y Apple, 1997; Griffiths, 2003; Kumashiro, 2015).

En esta misma línea, Freire (1970) apunta que el papel de la educación en la sociedad es la transformación. Combinar el lenguaje de la crítica con el lenguaje de la posibilidad, una combinación de utopía y práctica.

De acuerdo con Murillo y Hernández-Castilla (2014), y de acuerdo a esta doble visión de la escuela como reproductora o transformadora de las desigualdades sociales, “no todas las escuelas son iguales: algunas contribuyen al mantenimiento de las desigualdades, y otras hacen aportaciones significativas a la promoción de la Justicia Social” (p. 15). En primer lugar, el hecho de que las escuelas sigan reproduciendo las desigualdades existentes en la sociedad o bien luchen para cambiar dicha situación y promover una escuela y una sociedad más justa no es un elemento puramente estructural, sino de la conciencia e intencionalidad de la propia escuela.

En segundo lugar, consideramos que solo cuando el centro educativo hace explícito su intención de luchar contra las desigualdades e injusticias de la sociedad podrá lograrlo (Murillo y Hernández-Castilla, 2014). Si la escuela no se plantea esta lucha activa contra las inequidades, estará reproduciendo y manteniendo las desigualdades en la sociedad, convirtiéndose en un mecanismo del sistema para mantener las diferencias sociales. Los profesores Murillo y Hernández-Castilla (2014) defienden que es fundamental:

pensar en la necesidad de que la lucha de las desigualdades y a favor de una mayor Justicia Social debería formar parte de los proyectos educativos, de los objetivos que se plantean y, por supuesto, de las acciones que se llevan a cabo. (p. 17)

De esta forma, señalan que el elemento sustancial que puede lograr esta lucha real contra las desigualdades, es que el centro

escolar tenga una cultura para la Justicia Social. En este sentido, afirman que:

una escuela luchará para la Justicia Social en la medida que tenga una cultura compartida por la comunidad escolar, reflejada en su visión de la escuela y en sus acciones y comportamientos cotidianos, de lucha por las desigualdades, a favor de la equidad escolar y social. (Murillo y Hernández-Castilla, 2014, p. 17)

Uno de los estudios de referencia sobre la educación para la Justicia Social es el de Carlisle, Jackson y Alison (2006). El mismo se enmarca en un proyecto de Educación para la Justicia Social en las Escuelas (*SJES-Social Justice Education in Schools*) que busca impactar positivamente en la práctica de la Justicia Social en los entornos urbanos. Los autores se proponen identificar cinco principios clave de una educación para la Justicia Social basándose en el trabajo de una escuela cuyo alumnado está compuesto principalmente por estudiantes de raza negra, latinos, y/o pobres. Los cinco principios de la Educación en Justicia Social que identifican los autores son: inclusión y equidad (aborda todas las formas de opresión social como el racismo, el sexismo, la homofobia, el clasismo, etc.), altas expectativas hacia los estudiantes (la escuela proporciona un ambiente que confía en el desarrollo de todos los alumnos y alumnas), relaciones recíprocas con la comunidad (se busca activamente mantener una relación mutuamente beneficiosa con la comunidad-familias o agencias locales en un sistema abierto y accesible) un enfoque de todo el sistema (existe un compromiso de crear un ambiente socialmente justo dentro de todas las áreas del sistema) y compromiso con una educación liberadora (proporciona estrategias para intervenir en situaciones de opresión y busca facilitar el desarrollo de un pensamiento liberador). Los cinco principios identificados por

los autores ayudan a comprender cómo es la cultura de los centros que trabajan para la Justicia Social.

Complementando el anterior estudio, Maurianne Adams, Lee Anne Bell y Diane Griffin (2007) consideran que existen cinco principios básicos para una educación para la Justicia Social: a) establecer un equilibrio entre los componentes emocionales y cognitivos del proceso de aprendizaje de los estudiantes, b) reconocer y apoyar a las características personales e individuales de los alumnos y alumnas así como favorecer las interacciones sociales, c) prestar atención explícita a las relaciones sociales dentro del aula, d) favorecer la reflexión y la conciencia crítica de los estudiantes, y e) promover el esfuerzo y el crecimiento personal como base para el desarrollo integral del alumnado.

Desde la perspectiva de Connie E. North (2006), la Justicia Social en educación solo puede comprenderse a través de tres binomios claves: el binomio igualdad-diferencia, el binomio macro-micro y el binomio redistribución-reconocimiento. El primero hace referencia a uno de los debates más importantes en torno a una educación socialmente justa, centrado en tratar por igual a todos los estudiantes o bien de forma equitativa, es decir, adaptando la educación a las características y necesidades de cada uno de ellos. Así, es fundamental posicionarse en las opciones que individualizan y adaptan la enseñanza, respetando y favoreciendo la diferencia del alumnado. El segundo implica variar la atención en la educación, tanto en procesos de nivel macro como la formulación de políticas educativas y la organización de movimientos sociales, como en los procesos de nivel micro como los comportamientos individuales y las interacciones sociales desarrolladas en el aula. El tercer binomio ahonda en la concepción de la Justicia Social en

educación más cercana a la redistribución de los bienes primarios o bien en el reconocimiento de las particularidades culturales y étnicas de los alumnos. La autora defiende una posición dualista que conciba la educación para la Justicia Social como aquella que redistribuye los recursos y vela por reconocer y construir la identidad de sus estudiantes.

Estas pinceladas teóricas nos ayudan a esbozar la clara orientación social de la educación y su relevante función en la consecución de una sociedad más justa.

4.2. EVALUACIÓN DE ESTUDIANTES

La evaluación ha vivido profundos cambios en las últimas décadas, lo que ha complejizado la forma de entenderla así como sus implicaciones. El concepto de evaluación se ha ido abordando desde diferentes perspectivas añadiendo las necesidades y reflexiones del siglo XXI (Escudero, 2003; Lukas y Santiago, 2004; Sephard, 2000, 2006). La evolución del concepto ha permitido aclarar su significado, así como discernir entre evaluación y otros conceptos relacionados (Nevo, 1983).

Tan importante es la evaluación en el contexto actual que autores como Davidson (2004) consideran que la evaluación es la actividad más importante que nos ha permitido evolucionar, desarrollar, mejorar y sobrevivir en un entorno en constante cambio, especialmente dentro del campo social y educativo.

Considerando la importancia de la evaluación, es fundamental comprender de forma analítica el concepto de evaluación, desde concepciones más generales hasta la evaluación centrada en el ámbito educativo y en concreto la evaluación de los estudiantes.

4.2.1. Delimitación conceptual de la evaluación

El concepto de evaluación, lejos de tener una definición cerrada y definitiva está en constante evolución, siendo abierto y dinámico. Quizá, una de los motivos por el que está en constante conformación es porque ha recibido aportaciones de diferentes disciplinas y metodologías, estando arraigado a debates muy reñidos entre la filosofía de la ciencia y las teorías del conocimiento (Descy y Tessaring, 2004). Así, numerosos autores plantean que el concepto de evaluación nace en el campo empresarial y se extiende al resto de áreas del conocimiento (Casanova, 1998; Popham, 1990; Trochim, 2009). Desde esta perspectiva, la forma en que se defina la evaluación tiene altas implicaciones políticas, sociales y epistemológicas (Popham, 1990).

En sus inicios, la evaluación se concebía de una forma más estática, centrada en la valoración de los resultados finales de cualquier proceso y en la emisión de juicios de valor. Esta forma de comprender la evaluación centrada en los resultados o emisión de juicios de valor pretende medir los cambios inmediatamente después de la aplicación de una actividad humana o programa, estableciendo en qué medida estos cambios se han producido en respuesta a la intervención que se está evaluando (ICAP, 1995; Weiss y Rein, 1970).

La lógica de evaluación centrada en el producto o valoración final se expresa en una secuencia de cuatro etapas: a) el establecimiento de criterios de evaluación y las correspondientes dimensiones, b) la construcción de estándares de desempeño en relación a estos criterios y dimensiones, c) el rendimiento de medición en la práctica, y d) el alcance un

conclusión acerca del valor del objeto en cuestión (Scriven, 1991).

Esta primera aproximación al concepto de evaluación dio a paso a una concepción más centrada en el proceso y no tan solo en el resultado. Rossi, Lipsey y Freeman (2004) definen la evaluación como el proceso de determinar el crecimiento o fijar un valor para un objeto concreto, enfatizando en que durante el proceso de evaluación se recogen informaciones y evidencias que permiten no tan solo emitir un juicio de valor, sino también y especialmente entender dicho proceso. Complementariamente, Stufflebeam y Shinkfield (1987) la entienden como:

El proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y el mérito de las metas, la planificación, la realización y el impacto de un objeto determinado, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados. (p. 183)

Finalmente, encontramos una concepción de evaluación centrada no solo en el proceso y los resultados, sino también en decisiones que se tomen a partir de la información recopilada y la mejora que implica. Los conceptos planteados desde esta perspectiva plantean la evaluación desde su forma más compleja y difícil de llevar a cabo; y se basan en examinar los efectos a largo plazo de una intervención y su evaluación sobre los participantes (Alkin, 1969; Escudero, 2003; Guba y Lincoln, 1989; Madaus, 1991; Stufflebeam y Shinkfield, 1987).

Focalizando la evaluación en la toma de decisiones, la evaluación se define como el proceso de reconocer las áreas de decisión importantes, seleccionar la información apropiada y reconocer y analizar la información para transmitir datos que ayuden a tomar decisiones (Alkin, 1969; Cronbach, 1963).

Siguiendo esta definición, Tenbrink (1974, p. 19) describe “la evaluación como el proceso de obtención de información y su uso para formular juicios que a su vez se utilizarán en la toma de decisiones. Stufflebeam y Shinkfield (1987), amplían el concepto concretando que es el:

Proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y el mérito de las metas, la planificación, la realización y el impacto de un objeto determinado, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados. (p. 193)

Aunque podemos distinguir estas tres concepciones en la mayoría de definiciones de evaluación, no existe una acepción definitiva, por lo que comprender las diferentes perspectivas ayuda a dimensionar magnitud al concepto. Asimismo, las definiciones de evaluación más actuales suelen contemplar los tres elementos distinguidos en el análisis de las definiciones: el juicio de valor, el proceso y la toma de decisiones o mejora. En este sentido, Patton (2003) afirma que no siente ninguna necesidad de una definición definitiva basada en la autoridad de evaluación, y no espera que emerja una definición a pesar de la inseguridad que muchas personas plantean frente a dicha diversidad conceptual.

4.2.2. Aproximaciones a la evaluación en el ámbito educativo

Centrando la evaluación al campo de la Educación, Taras (2005) afirma que hay una falta de uniformidad en la definición de la terminología relativa a la evaluación educativa. No se trata de un concepto hermético sino que es polisémico (Olmos, 2008), precisamente por sus implicaciones sociales, éticas,

técnicas y políticas (Coll, 1999; Santos Guerra, 2003). Además, la profesora María Antonia Casanova (1998) añade otro de los problemas agregados al definir la evaluación educativa. Para la profesora, actualmente se hace patente una divergencia entre los conceptos de evaluación que se manejan a nivel teórico y la práctica real en las aulas.

Teniendo en cuenta estas consideraciones previas, a priori, la evaluación educativa se define como la aplicación de la evaluación en el ámbito educativo. Sin embargo, tiene una serie de características que la constituyen como un campo de estudio en sí mismo.

Uno de los profesores de referencia en la delimitación terminológica de la evaluación educativa fue Ralph Tyler (1951) considerado por muchos investigadores como el padre de la evaluación educativa (Escudero, 2003; García Garduño, 2005; Stufflebeam y Shinkfield, 1995, entre otros). Tyler (1951) define la evaluación educativa como el proceso de emisión de un juicio para los objetivos educativos (basándose en comportamientos observables) realizado a través de la educación y las actividades de aula. Inspirándose en la concepción tyleriana, Daniel Stufflebeam (1971) la concibe como el proceso de definir, adquirir y proporcionar la información necesaria para proceso de toma de decisiones en el ámbito educativo, y Nevo (1983) como una descripción sistemática de los objetos educativos y/o una valoración de su mérito o valor (Nevo, 1983; Popham, 1980). Taba (1962) concibe la evaluación en educación como:

El proceso intrincado y complejo que comienza con la formulación de objetivos, que involucra decisiones sobre los medios para asegurar la evidencia de su cumplimiento, los procesos de interpretación para llegar al significado de esta

evidencia y los juicios sobre las seguridades y las deficiencias de los estudiantes y que finaliza con las decisiones acerca de los cambios y las mejoras que necesitan el currículum y la enseñanza. (p. 409)

Siguiendo esta propuesta, Doval (1995, p. 213) define evaluación como “el proceso o resultado de la indagación de una realidad educativa a examen, que permite describirla, compararla con un estándar y emitir juicios ordenados a la toma de decisiones”.

Otro de los profesores referentes en la conceptualización de la evaluación en el campo educativo es Lee Joseph Cronbach y sus compañeros de investigación de la Universidad de Stanford. Estos investigadores comprenden la evaluación “como un examen sistemático de los acontecimientos que ocurren en y consecuencia de un programa-examen realizado para ayudar en la mejora de este programa y otros programas que tienen el mismo propósito general” (Cronbach, Ambron, Dornbusch, Hess, Hornik y Phillips, 1980, p. 14). Cronbach y sus colegas (1980) rechazan claramente el carácter arbitrario de la evaluación, ya que conciben al evaluador como un educador que debe juzgar lo que los estudiantes aprenden, en lugar de ser un árbitro que decide lo que es correcto o incorrecto.

Black y Wiliam (1998) y Sadler (1998) matizan que la evaluación educativa es el proceso mecánico o pasos para la emisión de un juicio del valor. Añaden, que las decisiones no pueden ser tomadas en el vacío, y por lo tanto, es necesario establecer puntos de comparación, normas y objetivos. Los criterios deben limitar las opciones que se consideren necesarias y relevantes para un juicio específico en el contexto educativo (Rodríguez-Gómez e Ibarra-Sáiz, 2012). Los autores, defienden que dentro del proceso se manifiestan los parámetros

dentro de los cuales se realiza el juicio. En otras palabras, durante el proceso de hacer un juicio, todos estos elementos están en constante interrelación. Todas las evaluaciones requieren estos parámetros, pudiendo ser explícitos o implícitos. Eisner (1985), por su parte, completa que se trata de un proceso que, en parte, ayuda a determinar si lo que hacemos en las escuelas está contribuyendo a conseguir los fines valiosos o si es antiético a estos fines.

Para Stufflebeam y Shinkfield (1995, pp. 26-27) la evaluación educativa tiene que cumplir las siguientes condiciones:

- Ser útil, facilitando informaciones sobre de virtudes y defectos así como soluciones para mejorar.
- Ser factible, empleando procedimientos evaluativos que se puedan utilizar sin mucho problema.
- Ser ética, basándose en compromisos explícitos que aseguren la necesidad de cooperación, la protección de los derechos de las partes implicadas y la honradez de los resultados
- Ser exacta, describiendo el objeto en su evolución y contexto, mostrando virtudes y defectos, al estar libre de influencias y al proporcionar conclusiones.

La evaluación educativa comprende la evaluación de programas, la evaluación de centros educativos, la evaluación de estudiantes y la evaluación de docentes, abarcando en sus definiciones diferentes objetos de la evaluación. En el punto siguiente, se va a abordar y analizar detenidamente el campo de la evaluación de estudiantes, que centra el objeto de la presente tesis doctoral.

Efectivamente, uno de los aspectos en los que la evaluación adquiere más importancia es en la valoración del recorrido de aprendizaje que cada alumno o alumna realiza. Así, determinar en qué medida el estudiante ha mejorado su aprendizaje o ha adquirido determinados objetivos es la herramienta fundamental del sistema educativo para determinar el rendimiento del alumnado (Sanmartí, 2007).

La evaluación de los estudiantes, usada en su acepción más amplia, incluye todos los métodos posibles para determinar el alcance de los estudiantes en cuanto a los objetivos educativos y los aprendizajes que se pretenden (Gronlund, 2003). Desde esta premisa, la evaluación del proceso aprendizaje es la reunión sistemática de evidencias, a fin de determinar si en realidad se producen ciertos cambios (aprendizajes) en los alumnos o no (Chadwick y Rivera, 1990).

El profesor Robert Stake (1967) concibe la evaluación de estudiantes como el proceso sistemático de emisión de juicios de valor basado en una descripción de las discrepancias entre las ejecuciones de estudiantes y profesores y los estándares del aprendizaje y de la enseñanza previamente definidos. En esta misma línea, Bloom (1975, p. 23) complementa esta visión afirmando que se trata de la “reunión sistemática de evidencias a fin de determinar si en realidad se producen ciertos cambios en los alumnos y establecer el grado de cambio en cada estudiante”. Gronlund (1985, p. 5), por su parte, la percibe como el “proceso sistemático de recopilar, analizar e interpretar información para determinar el grado en que los alumnos han logrado los objetivos instructivos”.

Con un enfoque más constructivista, Chadwick y Rivera (1990, p. 38) la definen como la “reunión sistemática de evidencia, a fin

de determinar si en realidad se producen ciertos cambios (aprendizajes) en los alumnos, y controlar, también, el estadio del aprendizaje en cada estudiante”.

De forma más centrada en el intercambio de información para la mejora, la evaluación educativa tiene por objeto determinar si los estudiantes están aprendiendo y se perfila como una parte integral de la búsqueda de la mejora de la educación. En este sentido, proporciona información a los estudiantes, los educadores, los padres, los responsables políticos y al público acerca de la eficacia de los servicios educativos (Pellegrino, Chudowsky y Glaser, 2001).

La evaluación de los estudiantes es un elemento fundamental para la mejora de la calidad de la educación, indicando que la retroalimentación es crucial para ayudar a los alumnos y alumnas a aprender. La evaluación debe reflejar una buena enseñanza por parte del docente y proporcionar información sobre los niveles de comprensión que los alumnos van alcanzando. A fin de que los estudiantes profundicen en su aprendizaje y su comprensión, una retroalimentación frecuente es fundamental: estos necesitan controlar y mejorar su proceso de aprendizaje así como autoevaluar activamente cómo aprenden (Bransford, Brown y Cocking, 1999; Rodríguez-Gómez e Ibarra-Sáiz, 2015).

Huba y Freed (2000) amplían que un ingrediente clave de la evaluación es que ésta permita cometer errores en el proceso de aprendizaje y aprender de ellos. Para ello, es esencial que el docente entrelace enseñanza y evaluación, dejando de ser prácticas episódicas separadas, sino más bien, estando en el mismo curso de acción, siendo actividades interrelacionadas

centradas en proporcionar orientación para la mejora tanto a los estudiantes como a los profesores y profesoras.

4.3. CONCEPCIONES SOBRE EVALUACIÓN DE ESTUDIANTES

Un elemento clave que determina las prácticas de evaluación son las concepciones que de las mismas tienen los diferentes agentes implicados (Coll y Remesal, 2009; Griffiths, Gore y Ladwig, 2006; Harris y Brown, 2009; Hidalgo y Murillo, 2017; Pope, Green, Johnson y Mitchell, 2009; Remesal, 2011; Quesada-Serra, Rodríguez-Gómez e Ibarra-Sáiz, 2016). De esta forma, es importante conocer qué son las concepciones, cómo se construyen y qué evidencias empíricas existen acerca de cómo estas influyen en la evaluación del aprendizaje de los alumnos y alumnas.

El apartado teórico de las concepciones coincide con uno de los artículos publicados que se escribió con el fin de realizar un análisis del estado de las investigaciones sobre las concepciones de evaluación de estudiantes. Por ese motivo, presentamos a continuación la información más relevante de la publicación.

4.3.1. Definición del término “concepciones”: del ámbito psicológico al educativo

Desde la perspectiva de la psicología cognitiva se entienden las concepciones como un elemento esencial para comprender la psique y el comportamiento humano, ya que hacen referencia a las representaciones de las personas del mundo que les rodea (Pozo, Scheuer, Mateos y Pérez-Echeverría, 2006). En este sentido, las concepciones se definen como un sistema organizado de creencias, originadas en las experiencias del individuo y desarrolladas a través de las interacciones en las

que participa (Coll y Remesal, 2009; Hasselgren y Beach, 1997; Martín, Mateos, Martínez, Cervi, Pecharromán y Villalón, 2006; Marton, 1981, 1986; Remesal, 2011; Säljö, 1994).

Las personas, por naturaleza, construimos nuestro pensamiento a partir de ideas, opiniones y creencias originada en nuestras propias experiencias. En función de las vivencias, cada ser humano concibe y construye en su mente el mundo de forma distinta al resto de personas, y esto incide en la forma en que se comporta e interacciona a lo largo de su vida. Por lo tanto, aquellas concepciones que tenga cada persona determinan en gran medida la forma en que actúa y la intencionalidad en dichos actos (Atkinson y Claxton, 2000; Harris, 2008; Murillo, Hidalgo y Flores, 2016; Pozo, 2006).

Una de las características principales de las concepciones es su carácter eminentemente social. Así, lejos de ser individuales, las concepciones nacen y se construyen en entornos sociales, ya que se fundamentan e interrelacionan en espacios y situaciones culturalmente compartidos. En este sentido, nuestras concepciones se construyen en la interacción con otras personas, por lo que la forma en que nosotros concibamos el mundo está influida directamente por los individuos que nos rodean (Pozo, 2006; Van den Berg, 2002).

Centrándonos en el ámbito educativo, las concepciones de los profesores y profesoras son una red de creencias, ideas y opiniones que influyen directamente en la forma en que estos entienden el proceso de enseñanza-aprendizaje e interactúan diariamente con sus estudiantes y el resto de compañeros docentes (Griffiths, Gore y Ladwig, 2006; Prieto y Contreras, 2008; Thompson, 1992). Las concepciones docentes se entienden, por tanto, como un marco o estructura de ideas

implícitas personales a través de las cuales los docentes interpretan su práctica profesional (Philipp, 2007; Pratt, 1992; Thompson, 1992). De esta forma, las concepciones del profesorado son influidas por el contexto en el que se desarrolla el profesor, definiendo su práctica docente. Queda claro que dicha influencia no es siempre coherente o explícita, pero sin duda las concepciones son un factor decisivo que determina el proceso de enseñanza-aprendizaje (Ekeblad y Bond, 1994; Griffiths, Gore y Ladwig, 2006).

Si trasladamos el concepto de concepciones a la evaluación, se entienden como aquellas ideas previas, creencias, teorías implícitas y estructuras mentales que dibujan la forma en la que los profesores y profesoras entienden la evaluación de los estudiantes así como su ejecución (Brown, 2004; Scott et al., 2014). La evaluación del alumnado puede ser concebida de distintas formas, existiendo múltiples concepciones de evaluación, las cuales, al tener un carácter social, suelen ser influidas por muchos factores, incluyendo el contexto y la cultura en la que trabajan, su formación previa o sus experiencias escolares vividas (Brown, 2003; Rubie-Davies, 2015; Scott et al., 2014). En este sentido, la forma en que evalúa un docente refleja, en parte, sus ideas implícitas construidas en interacción con otros sujetos sobre cómo entiende la evaluación, impactando directamente en su práctica evaluativa (Marshall y Drummond, 2006).

4.3.2. Revisión de las investigaciones centradas en las concepciones sobre evaluación de estudiantes

Un fructífera y reciente línea de investigación es la que estudia las concepciones sobre el proceso de evaluación del aprendizaje (Atjonen, 2014; Azis, 2012; Brown, 2003, 2004; Brown y

Hirschfeld, 2008; Brown y Remesal, 2012; Brown, Lake y Matters, 2011; Chen, Morris y Mansour, 2015; Fives y Gill, 2014; Harris y Brown, 2008, 2009; Winterbottom et al., 2008; entre otros).

Uno de los primeros estudios que se centró en las concepciones sobre evaluación fue el realizado por el profesor estadounidense M. Frank Pajares. En su canónico estudio, Pajares (1992) confirmó que profundizar en las concepciones de la enseñanza es prioritario para la investigación en educación, ya que sólo si conocemos las concepciones podemos cambiar las prácticas del profesorado en el aula. En este sentido, conociendo las creencias implícitas podremos determinar las prácticas docentes y así trazar líneas para su mejora, fin último de cualquier investigación educativa (Brown, 2008; Opre, 2010, 2015; Pajares, 1992; Sambell, McDowell y Brown, 1997; Vandeyar y Killen, 2007).

El interés evidente mostrado por la investigación en las concepciones sobre evaluación parte de la premisa de que estas concepciones, al igual que otros tipos de creencias de los profesores, influyen significativamente en sus decisiones y actividad profesional (Brown, 2008, Opre, 2010; Vandeyar y Killen, 2007). Por lo tanto, si queremos comprender cómo evalúan los profesores es fundamental conocer sus concepciones. Con esta lógica de fondo, nos interesa conocer las distintas direcciones que han cogido las investigaciones centradas en las concepciones sobre evaluación para aprender de ellas.

Los primeros estudios en los que vamos a profundizar son aquellos que persiguen conocer las concepciones sobre evaluación. Entre ellas, cobra especial relevancia el trabajo

realizado en la última década por Gavin T. L. Brown y sus colaboradores en Nueva Zelanda, el cual ha inspirado investigaciones posteriores con el mismo objeto de estudio (Brown, 2008, 2011; Brown y Harris, 2012; Brown y Hirschfeld, 2008; Brown, Irving, Peterson y Hirschfeld, 2009; Brown, Kennedy, Fok, Chan y Yu, 2009; Harris, Brown y Harnett, 2014; Hirschfeld y Brown, 2000).

En una serie de estudios iniciados por Brown y su equipo se propusieron conocer las concepciones acerca de la evaluación de docentes de Educación Primaria y Secundaria. El enfoque de su investigación fue cuantitativo y el análisis de los datos se realizó a través de modelos de ecuaciones estructurales. Para la recogida de datos se utilizó el inventario “*Teachers’ Conceptions of Assessment questionnaire (TCoA-III)*”. Los hallazgos de su investigación confirman que existen cuatro grandes concepciones que los docentes tienen sobre la evaluación del aprendizaje del alumnado (Brown, 2003, 2004, 2006; Brown y Hirschfeld, 2008; Harris y Brown, 2008, 2009):

- *La evaluación como mejora*, comprendida como una actividad que proporciona información útil para transformar la práctica docente así como el aprendizaje de los estudiantes.
- *La evaluación como herramienta de rendición de cuentas de la escuela*; es decir, como un instrumento para que los docentes demuestren que realizan correctamente su trabajo y que este sirva para que los estudiantes alcancen los estándares educativos planteados y sea, por lo tanto, de calidad.
- *La evaluación como instrumento de rendición de cuentas del propio estudiante* que demuestre lo que han aprendido y

hasta qué punto han alcanzado los objetivos de aprendizaje.

- *La evaluación como un proceso irrelevante*, concibiendo la evaluación como un proceso poco útil para el aprendizaje ya que solo sirve para calificar y clasificar al alumnado, lo cual afecta a su autoconcepto y no ayuda a su aprendizaje.

Profundizando en cómo se relacionan estas cuatro concepciones halladas en sus estudios, Brown (2003, 2004, 2006, 2009) determina que los docentes que tienen una concepción cercana a la evaluación como mejora están también moderadamente de acuerdo con la creencia de que la evaluación es una herramienta de rendición de la escuela ($r_{xy}=0,58$). Por otro lado, los profesores y profesoras que muestran menos cercanía con que la evaluación es irrelevante son los que muestran un acuerdo más elevado con la evaluación como mejora ($r_{xy}=-0,69$). Las concepciones de rendición de cuentas tanto de la escuela como de los propios estudiantes están moderadamente correlacionadas también ($r_{xy}=0,58$). Estos resultados muestran que los docentes o estudiantes no se ubican únicamente en una concepción de evaluación sino que, aunque suelen mostrar más acuerdo con una de ellas, sus creencias también tienen elementos o características del resto de concepciones. La investigación también revela que estas concepciones se mantienen constantes a lo largo del tiempo (Harris y Brown, 2008).

En el ámbito español, Ana Remesal (2011), de la Universidad de Barcelona, investigó las concepciones docentes de las funciones de la práctica evaluativa con un estudio cualitativo a través de entrevistas a treinta profesores. Los hallazgos

principales del estudio son que las concepciones de evaluación se encuentran en un continuo con dos polos: un polo pedagógico y un polo social. Los docentes que se encuentran en el extremo de la concepción pedagógica de la evaluación consideran que esta regula y ayuda al aprendizaje del alumnado mientras que los que se sitúan en el polo social defienden que la práctica evaluativa es un instrumento de acreditación social que sirve para clasificar y calificar a los estudiantes, no para aportarles información sobre cómo mejorar su aprendizaje. Frente a esta concepción dual de la evaluación, los resultados del estudio de Remesal (2011) reflejan que casi la mitad de los docentes tienen una concepción mixta más cercana al polo pedagógico, un 16% una creencia puramente pedagógica y otro 16% una concepción social. La autora concluye con que en general los docentes siguen mostrando concepciones arraigadas a las prácticas más tradicionales de evaluación.

Unos años más tarde, el investigador tunecino Sahbi Hidri (2015), en una investigación similar a la de Brown, se planteó conocer las concepciones de evaluación de los docentes de educación Secundaria y de Universidad. Aplicando el “*Teachers’ Conceptions of Assessment questionnaire* (TCoA-III)” junto con un cuestionario complementario adecuado al contexto de Túnez, Hidri (2015) utilizó el análisis factorial para analizar los datos. Los principales hallazgos son que el modelo de concepciones resultante estaba conformado por tres creencias de evaluación, concretamente que la evaluación sirve para la mejora, para la rendición de cuentas o que es irrelevante. Las concepciones de los docentes de mejora y rendición de cuentas correlacionaron positivamente. Estos hallazgos difieren en cierta medida por las investigaciones previas realizadas tanto por Brown como por otros investigadores. El autor concluye que estas diferencias

pueden deberse a la escasa formación en evaluación que recibe el profesorado y la mala reputación de la evaluación en su país, pudiendo incidir en las concepciones del profesorado sobre las prácticas evaluativas.

Con el propósito de conocer las concepciones de evaluación de los profesores de Secundaria y relacionarlas con su práctica profesional, el profesor de la Universidad de Wellington Astuti Azis (2014, 2015) realizó un estudio mixto en el contexto de Indonesia. Teniendo en cuenta las concepciones encontradas en la investigación de Brown (2004, 2006, 2009), los participantes de este estudio consideran que la evaluación a) sirve para mejorar la enseñanza y el aprendizaje, b) sirve para la rendición de cuentas de la escuela, y c) es útil para que los estudiantes demuestren aquello que han aprendido. Por otro lado, los resultados muestran que los docentes no están de acuerdo con que la evaluación sea irrelevante. La parte del estudio referente a la vinculación de las creencias con sus prácticas, las mostramos en el epígrafe dedicado a este aspecto del presente artículo.

Otro aspecto en el que se centran estos estudios es no tan solo en las concepciones de docentes, sino también en las de los estudiantes. Con este fin, Brown y Hirschfeld (2008) realizaron un estudio con 3.469 alumnos de Educación Secundaria con el objetivo de determinar las concepciones de los alumnos acerca de la evaluación. El análisis de ecuaciones estructurales arrojó que los estudiantes tienen cuatro tipos diferentes de concepciones: a) evaluación como mejora del rendimiento y del aprendizaje, b) la evaluación hace que los estudiantes rindan cuentas y sean responsables, c) la evaluación es irrelevante, y d) la evaluación es agradable. Los autores concluyen que la primera concepción correlaciona positivamente con el

rendimiento, mientras que los estudiantes con las concepciones c y d son los que peor desempeño obtienen.

Estos resultados nos permiten hallar similitudes entre las concepciones de docentes y estudiantes. Así, tres de las creencias son iguales para ambos mientras que hay una concepción diferente (para los docentes la evaluación como rendición de cuenta de la escuela frente a que la evaluación es agradable para el alumnado). En este sentido, existe coherencia y estabilidad también en las concepciones del alumnado, del mismo modo que sucede con las de los profesores y profesoras (Brown, 2003, 2004; Brown y Hirschfeld, 2008; Harris y Brown, 2008).

En un estudio posterior, Brown, Irving, Peterson y Hirschfeld (2009) matizaron que las concepciones de los estudiantes están más cercanas a la evaluación como una práctica formativa y no de control por parte del docente. Poco después, Brown y Harris (2012) complementaron dichos resultados llevando a cabo un nuevo estudio con el objetivo de conocer si existían diferencias en las concepciones de los estudiantes en función de su nivel de escolarización. Los resultados muestran que los estudiantes de educación secundaria, frente a los de primaria, puntúan más alto en concebir la evaluación como irrelevante, frente a concebirla como externa, como herramienta de mejora o como agradable. No obstante, en general, todos los estudiantes, sin diferencias significativas, conciben la evaluación en mayor medida como una herramienta que ayuda a mejorar su aprendizaje.

En Educación Secundaria, Peterson e Irving (2008) indagaron sobre las concepciones que tienen los estudiantes sobre la evaluación y su comunicación. El estudio, de corte cualitativo,

llevó a cabo cinco grupos focales en cuatro escuelas de Nueva Zelanda. Los resultados señalan que los estudiantes conciben la evaluación como útil cuando esta sirve para mejorar su aprendizaje e irrelevante cuando no proporciona información útil para ellos. Los autores concluyen que su estudio confirma de forma cualitativa los hallazgos de los estudios cuantitativos anteriores realizados por Brown y colaboradores.

En una línea similar de trabajo, Segers y Tillema (2011) realizaron una investigación en Países Bajos con 712 estudiantes con el objetivo de conocer sus concepciones acerca de la finalidad que tiene la evaluación. Los análisis factoriales realizados apuntan a que existen cinco concepciones sobre la finalidad de la evaluación: a) es beneficiosa para el aprendizaje, b) sirve como instrumento para medir lo que los estudiantes han aprendido, c) resulta agradable a los estudiantes, d) mejora el clima de apoyo y colaboración en el aula, y e) sirve para la rendición de cuentas. Los autores señalan que los estudiantes que perciben la evaluación como un instrumento para la rendición de cuentas la asocian a una práctica profundamente injusta.

Conocidas las concepciones de profesorado y alumnado, otro aspecto investigado por Brown y sus colaboradores, es comparar las concepciones de evaluación en distintos países. Así, en distintos estudios realizados se compararon las concepciones de profesores y estudiantes de Nueva Zelanda con los de Australia, España, Irán, Chipre y China. Los resultados por países permiten afirmar que existen diferencias entre las creencias de los participantes neozelandeses y del resto de países. No obstante, las concepciones de los docentes de los distintos países son un fiel reflejo del pensamiento de profesores y alumnos acerca de la naturaleza y el propósito de

la evaluación, la enseñanza y los planes de estudio de dichos estados. Las diferencias explicativas se deben a la diversidad cultural y de los sistemas educativos y son coherentes con las políticas educativas de los distintos países analizados. Como pinceladas de los resultados de cada uno de los estudios:

- En el caso de Queensland (Australia), Brown y Lake (2006) se realizó un estudio ex post facto y un análisis a través ecuaciones estructurales. Los resultados determinaron que existen similitudes en las concepciones de evaluación como mejora e irrelevante pero diferencias en las creencias sobre la evaluación como rendición de cuentas. Por su contexto de gestión y de evaluaciones externas, los docentes de Nueva Zelanda tenían una visión más cercana a la rendición de cuentas de los estudiantes a través de la evaluación, frente a un menor acuerdo por los profesores y profesoras de Queensland. En este caso, ambos grupos de profesores también muestran múltiples concepciones que correlacionan entre sí de forma estable.
- En el contexto español, el propósito del estudio fue comparar las concepciones de evaluación de docentes en formación neozelandeses ($n=324$) y españoles ($n=672$). Los autores consideran que es importante conocer las concepciones de evaluación de los estudiantes de magisterio, dado que tienen una amplia experiencia como estudiantes y que los modelos utilizados para el análisis de las concepciones con los docentes en ejercicio no sirven para dichos estudiantes. En este caso, los hallazgos confirmaron que existen diferencias estadísticas significativas en las concepciones de los

docentes en formación. Mientras que para los estudiantes de Nueva Zelanda, sus concepciones de la evaluación como mejora fueron inversas a la creencia de que la evaluación es irrelevante y positiva pero moderada con la rendición de cuentas, para los docentes en formación españoles las correlaciones fueron mucho más fuertes en ambos casos (Brown y Remesal, 2012). Como conclusión, los estudiantes de magisterio de España tienen una concepción sobre la evaluación más cercana a la mejora y la rendición de cuentas del propio estudiante.

- Con el propósito de conocer cómo conciben los profesores la evaluación en Chipre en comparación a Nueva Zelanda, Brown y Michaelides (2011) desarrolló un estudio *ex post facto* y posteriormente un análisis factorial exploratorio. Los resultados muestran que los docentes de Chipre tienen concepciones ligeramente diferentes a los docentes de Nueva Zelanda, debidas a las políticas y forma de llevarse a cabo la evaluación en ambos contextos. Las principales diferencias se encuentran en que para los profesores de Chipre las concepciones de la evaluación como mejora tienen una correlación moderada en considerar la evaluación como negativa o irrelevante y positiva con que la práctica evaluativa suponga una rendición de cuentas por partes de la escuela.
- Brown, Pishghadam y Sadafian (2014) analizaron si 760 estudiantes universitarios iraníes tienen una estructura de concepciones similar a los neozelandeses. Para ello, se hizo un análisis factorial confirmatorio y se obtuvo

que las concepciones que tenían diferían ligeramente, siendo: a) la evaluación mejora el aprendizaje y enseñanza, b) la evaluación sirve para que las escuelas y facultades rindan cuentas, y c) la evaluación tiene efectos negativos sobre el aprendizaje.

- En China, los profesores Gavin T. L. Brown y Lingbiao Gao (2015) examinaron cómo diversas investigaciones describen las concepciones y las prácticas de evaluación a través de un meta-análisis de ocho estudios de entrevistas y encuestas. En el análisis de las concepciones de evaluación de los docentes de China estos difieren al compararlos con los neozelandeses en considerar que la evaluación solo servirá para la mejora de la práctica docente y del aprendizaje del alumnado siempre y cuando cumpla su función ineludible de rendición de cuentas, tanto de la escuela como de los estudiantes.

Más allá de identificar las concepciones, otros investigadores buscaron comparar las creencias entre docentes y estudiantes universitarios (Fletcher, Meyer, Anderson, Johnston y Rees, 2012). Para ello, los autores realizaron un cuestionario para ambos grupos participantes y lo analizaron con Modelos de Ecuaciones Estructurales. Las diferencias en las concepciones de ambos grupos fueron claras. Mientras que los docentes universitarios concibieron la evaluación como un proceso que ayuda a mejorar la enseñanza y el aprendizaje, los estudiantes afirmaron que esta solo sirve para que ellos rindan cuentas, siendo irrelevante para su aprendizaje. Para los autores es fundamental reflexionar acerca de la importancia de que la evaluación sea un proceso compartido entre docentes y

estudiantes, sino, difícilmente podrá ayudar a los estudiantes y por ende, a la mejora del sistema educativo.

Las diferencias en las concepciones se originan en primer lugar por los participantes de la investigación, encontrando distintas concepciones si son estudiantes de educación obligatoria, futuros profesores o docentes en activo. En segundo lugar también encontramos diferencias por el nivel educativo, pudiendo estar las concepciones vinculadas a la evaluación que se desarrolla en Primaria, Secundaria o Educación Superior. En tercer lugar, las concepciones parecen ser estables en el tiempo y están relacionadas con el rendimiento en los estudiantes. Por último, las diferencias hemos visto que también pueden deberse al contexto del estudio, pudiendo afirmar que las concepciones se ven enormemente influidas por las políticas evaluativas del país o las directrices evaluativas del centro educativo. La influencia del contexto así como de otros factores se va a analizar en profundidad en el próximo apartado del presente artículo.

4.3.3. Incidencia del contexto en las concepciones sobre la evaluación de estudiantes

En uno de sus primeros estudios, Gavin T. L. Brown (2004) afirmó que, inicialmente, las concepciones no se ven determinadas por el contexto en el que se desarrollan o por experiencia previa. Para ser más precisos, el autor concluyó que el número de años en la educación, el número de años de experiencia profesional y la situación socioeconómica de sus escuelas no influye en las concepciones que los docentes tenían en relación con evaluación. Frente a esta afirmación, otros investigadores si han hallado diferencias en las concepciones de evaluación en función de distintos factores.

Efectivamente, el contexto es uno de los aspectos más estudiados por su relevancia en la conformación de las creencias de la evaluación del aprendizaje de los estudiantes (Rubie-Davies, Flint y McDonald, 2012; Vandeyar y Killen, 2007). La investigación ha evidenciado que las características sociales, culturales, económicas y políticas del país, ciudad o barrio donde trabajen los docentes o estudien los alumnos influyen directamente en las creencias implícitas que tengan de evaluación (Davis y Neitzel, 2010; James y Pedder, 2006; Prieto, 2008; Vandeyar y Killen, 2007).

Para la investigadora chilena Marcia Prieto (2008), las concepciones se ven influidas por el contexto en el que se desarrollan. Así, las creencias docentes están determinadas por el nivel socioeconómico, cultural y lingüístico de los estudiantes y del centro escolar. Los contenidos, los instrumentos y la finalidad de la evaluación planteada por el profesor cambian de una escuela a otra y de un estudiante u otro. Los profesores que trabajan en centros educativos ubicados en contextos desafiantes evaluarán teniendo en cuenta que sus estudiantes tienen una menor capacidad o están en desventaja social. En este sentido, la autora defiende que “la diversidad social y cultural que les caracterizan y por lo tanto, la evaluación se transforma en la vía que determina su trayectoria escolar, sobre la base de la calidad de su capital cultural” (Prieto, 2008, p. 134).

Vandeyar y Killen (2007), en una investigación de carácter etnográfico llevada a cabo en cuatro centros escolares de Sudáfrica, analizaron las creencias de sus docentes y concluyeron que el sistema dentro del cual los profesores y profesoras enseñan influye directamente en sus concepciones de evaluación. En este sentido, si el centro educativo hace hincapié

en el contenido, las imposiciones políticas en términos evaluativos y las pruebas de evaluación externa, no es sorprendente que los profesores tengan concepciones de evaluación próximas a que la función de esta es la rendición de cuentas del alumno o alumna y de la propia escuela. Por el contrario, si el centro educativo se preocupa del aprendizaje de sus estudiantes y de que la evaluación sea una herramienta de ayuda para estos, los docentes muestran concepciones más próximas a la mejora. En este sentido, la política, cultura y directrices evaluativas que sigan las escuelas determinan las concepciones de la práctica evaluativa de su profesorado.

No obstante, Vandeyar y Killen (2007) señalan que es muy iluso pensar que si cambian las políticas o cultura de evaluación de los escolares se transformaran automáticamente las concepciones de los docentes. La construcción de las creencias implícitas es un proceso complejo y por ello, el cambio en las concepciones también lo es. Ilustrando esta idea, el profesor británico Harry Torrance (1995) afirma que:

Creer que las concepciones de los docentes van a cambiar simplemente porque las políticas y los contextos escolares cambien es ingenuo, sin embargo, la capacidad y la voluntad de los docentes para implicarse con los cambios en el proceso de evaluación en particular, si se ve influida por sus ideas y creencias arraigadas. Por lo tanto, en un proceso de cambio evaluativo es fundamental implicar a los docentes en ella y que compartan su propósito, de lo contrario, no van a cambiar sus concepciones ni con ello su práctica evaluativa.
(p. 55)

Aportando más datos empíricos de las ideas presentadas sobre la influencia del contexto en las concepciones de evaluación de los profesores, el estudio de Brown, Hui, Yu y Kennedy (2011) es un referente al respecto. En su investigación realizada en China los autores destacan la incidencia del contexto en la

evaluación de los estudiantes, considerando que la cultura, así como el estatus social de los docentes incide en la forma en que estos conciben la enseñanza y la evaluación. Así, estos autores determinan que mientras que algunos factores pueden ser estables en toda la población, otros, como las concepciones de evaluación, pueden ser legítimamente divergentes por las diferencias sociales y culturales existentes en las escuelas y en las características sociales de estas.

En este sentido, la investigación fenomenográfica de Murillo, Hidalgo y Flores (2016) con docentes de educación primaria y secundaria determinó que el contexto donde estuviese ubicada la escuela donde el docente desarrolla su labor influye en el tipo de concepciones de evaluación que tenían. Así, los profesores que trabajan en contextos desafiantes muestran unas concepciones de evaluación más cercanas a la inclusión, adaptación e individualización de la evaluación, mientras que los docentes de escuelas situadas en entornos favorables tienen creencias más cercanas a la excelencia, la objetividad y la igualdad en la evaluación. En palabras de los propios autores:

El contexto determina las concepciones y con ello las prácticas. De esta forma, frente al supuesto tratamiento indiferenciado por parte de las Administraciones de todos los centros, la investigación muestra que la segregación escolar configura enseñanzas muy diferentes en función del contexto socio-económico. Si tenemos bien asumida la idea de que hay que dar un mayor apoyo al estudiante que más lo necesita, no hay razones para no extrapolar esa idea a los centros. [...] Al final, es el contexto el que está determinando el tipo de evaluación, de enseñanza y con ello el futuro de los niños y las niñas. (pp. 275-276)

4.4. CONCEPCIONES SOBRE UNA EVALUACIÓN JUSTA DE ESTUDIANTES

Una de las líneas de investigación que está cogiendo más fuerza entro del estudio de las concepciones sobre evaluación, es la investigación sobre las ideas implícitas de una evaluación justa. Esta línea de investigación aporta evidencias tanto teóricas como empíricas que son relevantes en la forma de comprender una evaluación más justa (Gipps y Stobart, 2009; Goodwin, 2012; Popham, 2004; Syverson, 2009).

Aunque la enseñanza y la evaluación han cambiado mucho en los últimos 50 años, las prácticas de evaluación siguen siendo injustas en gran medida en la actualidad. En palabras de Aitken (2012):

Gran parte de las viejas actitudes acerca de la clasificación y evaluación se están utilizando como armas para coaccionar, premiar y castigar a prevalecer. Si bien estas prácticas no están a la orden, algunas prácticas de evaluación injustas son todavía evidentes en la actualidad y siguen afectando a las vidas de demasiados estudiantes. (p. 179)

Definir qué es una evaluación justa no es una tarea fácil, ya que tiene implicaciones éticas que influyen enormemente en el aprendizaje de los estudiantes. Uno de los elementos que más dificultan la conceptualización de una evaluación justa es el binomio igualdad/equidad; es decir, una concepción más próxima a la igualdad o más cercana a la equidad.

Para profundizar en qué entendemos por una evaluación justa hemos organizado este apartado en dos epígrafes. Primero expondremos las distintas posturas teóricas acerca de cómo se conceptualiza y concibe una práctica evaluativa justa para, posteriormente, profundizar en las investigaciones empíricas que existen sobre esta temática.

4.4.1. Perspectivas teóricas de la evaluación justa de estudiantes

En este epígrafe mostramos las principales posturas teóricas relacionadas con una evaluación de estudiantes justa. Así, algunos autores resaltan aquellos aspectos más cercanos a la igualdad mientras que otros defienden que una evaluación justa es aquella que es equitativa. En general, existe una tendencia a que los autores expliciten cuales creen que son las características, procesos o requisitos necesarios para desarrollar una evaluación justa (Aitken; 2011; Cline, 2000; Darling-Hammond, 1994; Gilliland, 1993; Lam, 1995; Landy, Barnes-Farrell y Cleveland, 1980; Murillo e Hidalgo, en prensa; Picón-Jácome, 2013; Pettifor y Saklofske, 2012; Rogers, 2010; Suskie, 2002).

Desde la perspectiva de Landy, Barnes-Farrell y Cleveland (1980) existen varios procesos que favorecen una evaluación justa en términos de imparcialidad y precisión en la evaluación desempeñada y la calificación de los estudiantes. Estos procesos son: a) la frecuencia en la retroalimentación, b) la importancia de que los estudiantes conozcan los resultados de la evaluación, c) que la evaluación sea acorde con el aprendizaje desarrollado en el aula, y d) la proactividad de los docentes para ayudar a sus estudiantes a mejorar su aprendizaje y por ende su evaluación.

En la misma línea, Gilliland (1993) elabora un listado con lo que él entiende como las diez reglas básicas para desempeñar una evaluación justa:

1. Explicación clara y detallada de la evaluación.
2. Favorecer la comunicación entre los estudiantes y el docente.

3. Evaluación relacionada con los contenidos y procedimientos trabajados en el aula.
4. Consistencia de la prueba de evaluación.
5. Evitar el azar en las pruebas evaluativas.
6. Dar diferentes oportunidades a los estudiantes en la evaluación.
7. Uso de la retroalimentación.
8. Justificación y explicación de las decisiones tomadas a partir de la evaluación.
9. Consistencia de las pruebas de evaluación.
10. Apertura de la evaluación a los estudiantes a través de procedimientos como la coevaluación.

Por su parte, Lam (1995) indica que los docentes que conciben una evaluación justa desde una perspectiva igualitaria se centran en garantizar la igualdad de condiciones en la realización de la prueba (materiales, tiempo, espacio, recursos...). La evaluación igualitaria requiere la misma administración, contenidos, puntuación e interpretación de los resultados, persiguiendo, por lo tanto, la objetividad.

Frente a esta postura, autores como Suskie (2002) defienden que una evaluación justa no es aquella que es igual para todos los estudiantes, ya que no todos han tenido necesariamente las mismas oportunidades de acceso a la educación; por lo que el principio de igualdad no es suficiente para garantizar la justicia. Suskie (2002) considera esencial la promoción de la equidad en la evaluación para garantizar la justicia, adaptando la evaluación a las necesidades y características de los estudiantes. Esta equidad se traduce en llevar a cabo múltiples evaluaciones

y con diferentes instrumentos para asegurar que la diversidad de los alumnos y alumnas se atiende correctamente a través de la evaluación.

Complementando esta visión, Picón-Jácome (2013) considera que:

La justicia como un principio que implica el desarrollo de prácticas evaluativas que garanticen la equidad y que sean además válidas y transparentes. Considero asimismo que la búsqueda de la justicia debe incluir el desarrollo de formas de evaluación que sigan principios democráticos, es decir participativos e incluyentes, en la toma de decisiones dentro del aula. Estos principios y cualidades se tornarían a su vez en lineamientos a seguir en el diseño de los procedimientos de un sistema de evaluación que garantice un impacto positivo en los procesos de enseñanza y aprendizaje. Dicho impacto, que se traduce en una evaluación para el aprendizaje, debe ser el principal objetivo que guíe todo evento evaluativo dentro del aula. (p. 3)

Reflexionando acerca del binomio igualdad-equidad, Lam (1995) considera que garantizar la igualdad permite comparar los resultados y simplificar las evaluaciones realizadas, pero se pierde significación y puede no ser imparcial; mientras que garantizar la equidad reduce el sesgo y permite una evaluación significativa pero introduce dificultades en la administración y en la comparación de los resultados de los estudiantes.

En coherencia con los principios de igualdad y equidad, Lam (1995) reflexiona que una evaluación es injusta si 1) a los alumnos no se les brinda igualdad de oportunidades para demostrar lo que saben, por lo que las evaluaciones son parciales, 2) estas evaluaciones parciales se utilizan para juzgar las capacidades y necesidades de los alumnos y 3) si los docentes tienen una visión distorsionada del estudiante y esta

visión se traduce en la toma de decisiones educativas que finalmente limita las oportunidades educativas del alumno.

En cuanto a los cambios necesarios que la evaluación tiene que sufrir para ser justa, Darling-Hammond (1994) sostiene que: a) de plantearse como un mecanismo de clasificación a un soporte de diagnóstico; b) de ser una herramienta externa para medir el rendimiento a servir a los estudiantes y los docentes para profundizar en la enseñanza y el aprendizaje; y c) de servir para sancionar a aquellos que no cumplen los requisitos a servir para igualar la evaluación para todos los estudiantes.

En el ámbito británico, el investigador Tony Cline (2000) reflexiona acerca de la diversidad de estudiantes que existen en las escuelas de Reino Unido y considera que frente a esa realidad, es necesario desarrollar una evaluación justa para ayudar en el aprendizaje de todos y cada uno de los alumnos y alumnas. El autor defiende que una evaluación justa debe cumplir los principios de a) integridad teórica, teniendo en cuenta el desarrollo, motivaciones e intereses de los estudiantes, b) eficacia práctica, aportando información útil para docentes y estudiantes, c) equidad, adaptándose a los distintos alumnos y alumnas y sus características, d) rendición de cuentas referente a que las evaluaciones externas aporten una información comprensible para los estudiantes y sus familias y que ayude a mejorar en su aprendizaje, e) llevar a cabo una evaluación normativa, comparando a los estudiantes con sus iguales, especialmente con aquellos que tienen sus mismas características, y f) currículum relacionado con la evaluación, enseñando aquello que queremos que aprendan los estudiantes y con un nivel de dificultad adaptada a cada alumno o alumna.

Uno de los documentos referentes sobre una evaluación justa es el *The Principles for Fair Student Assessment Practices for Education in Canada*, elaborado por el *Joint Advisory Committee* en Canadá. Como miembro de dicho organismo, Todd Rogers (2010) desarrolló un conjunto de principios y directrices sobre una evaluación justa y equitativa. Esta guía de principios justos se envió a 32.000 profesionales del sector educativo de Canadá, los cuales han hecho sus aportaciones, críticas y sugerencias de mejora. Los principios finales se agrupan en dos apartados: el apartado A, referente a la evaluación de educación primaria y secundaria, y el apartado B, relacionado con la aplicación de pruebas estandarizadas. En el apartado A encontramos los siguientes principios:

- Usar métodos de evaluación variados y que estén relacionados con los objetivos de aprendizaje planteados. Es fundamental también que la evaluación sea continua, para tener una idea clara del progreso de cada estudiante.
- Los estudiantes deben ser informados de los criterios con los que serán evaluados, usando métodos cualitativos y favoreciendo que las instrucciones de la evaluación sean claras y concisas para la comprensión de los alumnos y alumnas.
- Los procedimientos para juzgar o calificar el rendimiento de los estudiantes deben ser apropiados para el método de evaluación utilizado y ser consistentemente aplicados así como monitoreados.

- Los resultados de la evaluación de cada estudiante deben ser detallados, precisos y reflejar el proceso de aprendizaje de cada estudiante.
- Es necesario ofrecer un informe individual a los estudiantes con el detalle de la evaluación de cada alumno o alumna, acompañado de una retroalimentación oral que ayude al estudiante a guiar su aprendizaje.

En el apartado B sobre las evaluaciones estandarizadas los principios hacen referencia a:

- Determinar claramente el propósito y finalidad de las evaluaciones realizadas así como qué se espera que sepan hacer los estudiantes en cada una de las preguntas o apartados de la prueba.
- Dar instrucciones claras a los estudiantes para realizar la prueba, de tal forma que no haya dudas en la realización de la misma. Asimismo, es fundamental que existan rúbricas claras y precisas sobre cómo se va a corregir la prueba y qué puntuación le pertenece a cada apartado.
- Promover una comunicación directa y fluida tanto con los estudiantes como con los padres o tutores de los estudiantes, con una información completa y presentada de forma comprensible.

Jean L. Pettifor y Donald H. Saklofske (2012), a la luz del documento *The Principles for Fair Student Assessment Practices for Education in Canada* (mencionados anteriormente), considerado por los autores como la guía más práctica para la

mejora de las prácticas de evaluación de estudiantes, proponen cinco principios que caracterizan la evaluación justa: a) desarrollo y selección de métodos de evaluación variados, claramente relacionados con los objetivos de aprendizaje de los estudiantes y adaptados a las necesidades de los estudiantes, b) recogida de información donde los estudiantes participen de forma activa y desarrollen procedimientos como la coevaluación y donde estos tengan claros y por escrito los criterios de desarrollo de esa prueba de evaluación, c) los criterios de corrección y calificación de las evaluaciones deben ser transparentes para los estudiantes, y los docentes deben devolver por escrito los resultados de evaluación de los estudiantes, con indicaciones claras en la mejora de su aprendizaje, d) evitar la interpretación sesgada de resultados por parte de los profesores, y e) devolver de forma periódica la información obtenida con la evaluación de forma clara, precisa y útil para los estudiantes.

Complementariamente, Pettifor y Saklofske (2012) afirman que existen diversas cuestiones pedagógicas que transforman la evaluación en una práctica más justa. Una de estas cuestiones es que los estudiantes conozcan los criterios de evaluación con los que van a ser evaluados, o incluso mejor, que dichos criterios sean co-definidos entre los docentes y los estudiantes (Andrade y Du, 2007; Harlen, 2005; Pepper y Pathak, 2008). Otro elemento importante para conseguir una evaluación más justa es no discriminar a los estudiantes por sus características sociales, culturales o económicas (que suelen tener impacto en la evaluación) (Stobart, 2005). Los autores también remarcan que los docentes no solamente evalúen las pruebas escritas sino también la personalidad, progreso, esfuerzo y participación de

los estudiantes, ya que solo añadiendo estos aspectos más cualitativos a la evaluación se logrará una práctica evaluativa justa. Asimismo, los autores destacan es esencial que la evaluación sea individualizada a las necesidades de cada estudiante.

En síntesis, podemos afirmar que la forma de entender la evaluación justa de los distintos investigadores analizados se encuentra más próxima a la equidad, entendiendo la justicia en términos evaluativos es aquella que se adapta a los estudiantes, se preocupa por su aprendizaje y progreso, utiliza diversidad de instrumentos (especialmente cualitativos) y devuelve la información a los estudiantes de forma constante para ayudar a la mejora del alumnado en su aprendizaje.

4.4.2. Principales estudios empíricos de una práctica evaluativa justa

Una vez realizada una exhaustiva búsqueda de los estudios acerca de qué es una evaluación justa, podemos afirmar que existen pocas investigaciones empíricas que arrojen datos sobre el tema. Para describir los estudios centrados en la evaluación justa los hemos organizado a partir de un criterio cronológico, ayudándonos a entender asimismo su evolución. Existen seis estudios referentes que profundizan en cómo tiene que ser la evaluación de estudiantes para ser justa (Aitken, 2012; Chilisa, 2000; Sambell, McDowell y Brown, 1997; Pepper y Pathak, 2008; Scott et al., 2014; Tierney, Simon y Charland, 2011). En la tabla 4.1 mostramos el detalle de estos estudios empíricos analizados.

A continuación exponemos los estudios que acabamos de mencionar. En cada uno de ellos explicamos, en primer lugar, el

objetivo y metodología utilizados y en segundo lugar destacamos sus principales hallazgos.

TABLA 4.1. RELACIÓN DE ESTUDIOS EMPÍRICOS SOBRE UNA EVALUACIÓN JUSTA

AUTORES	AÑO	PAÍS	OBJETIVO DE LA INVESTIGACIÓN
Sambell, McDowell y Brown	1997	Reino Unido	Conocer las concepciones de evaluación justa desde la perspectiva del alumnado.
Chilise	2000	Botswana	Examinar el lenguaje justo/injusto de las pruebas de evaluación.
Pepper y Pathak	2008	USA - Texas	Conocer las concepciones de una evaluación justa de los estudiantes y cómo influyen en las prácticas de evaluación que se desarrollan en el aula.
Tierney, Simon y Charland	2011	Canadá	Conocer las prácticas de evaluación justas llevadas a cabo por los docentes.
Aitken	2012	Canadá	Conocer las prácticas de evaluación justa que se desarrollan en el aula desde la perspectiva de los alumnos y alumnas.
Scott, Webber, Lupart, Aitken y Scott	2014	Canadá	Profundizar en las concepciones de docentes acerca de una evaluación justa.

Fuente: Elaboración propia.

El estudio realizado por Kay Sambell, Liz McDowell y Sally Brown (1997) en Reino Unido fue una de las primeras investigaciones en profundizar en cómo tiene que ser una evaluación justa. El objetivo del estudio fue determinar las percepciones de una evaluación justa desde la perspectiva de los estudiantes así como el impacto de la validez de dicha práctica evaluativa. La metodología utilizada fue el desarrollo de trece estudios de casos para profundizar en distintos métodos de evaluación alternativos y la validez de los mismos a través de entrevistas a estudiantes y docentes, la observación y la revisión de materiales relevantes para la evaluación. Los principales hallazgos del estudio muestran que: a) la evaluación

de corte tradicional contamina la evaluación alternativa, percibida como un proceso más justo para el aprendizaje de los estudiantes, b) consideran que la evaluación final, especialmente a través de un examen, es profundamente injusta y no refleja todo aquello que han adquirido y aprendido y c) los estudiantes destacan que una evaluación, para ser justa, debe cercana a sus vidas y ayudarles a desarrollar habilidades que ellos mismos utilicen en sus vidas diarias. Todo conocimiento poco útil para los alumnos y alumnas es percibido como injusto. Como conclusión final, los autores resaltan que los estudiantes no hacen referencia al término de validez al hablar de una evaluación justa, sino más bien a la imparcialidad, concepto más cercano a la equidad en la evaluación.

Tres años más tarde y focalizando la justicia evaluativa en cuanto al género, Bagele Chilisa (2000) realiza un análisis del papel del lenguaje en los exámenes de ciencias *Junior Certificate* que se aplican en Botswana. La muestra seleccionada estuvo compuesta por 42 escuelas y 8264 estudiantes (3999 hombres y 4265 mujeres) que habían realizado dicho examen. A través de un procedimiento estadístico de validación de los ítems de la prueba (MH-CHSQ), Chilisa identifica aquellos elementos en los que existe diferencia en cuanto a la respuesta de hombres y mujeres. Así, la autora afirma que los diferentes modos de evaluación denotan diferentes formas de conocimiento que no son necesariamente neutrales en cuanto al género en los exámenes de ciencias de *Junior Certificate*. La investigación concluye que existe una necesidad de revisión de las políticas de evaluación para que sean justas en cuanto al género, desarrollando pruebas que, en primer lugar, tengan contenidos que favorezcan a chicos y chicas y en segundo lugar que no tengan un lenguaje sexista.

Un tercer estudio referente en esta línea de investigación es el realizado por Pepper y Pathak (2008), quienes realizaron una investigación en la universidad de Southwestern (Texas) con el propósito de examinar las percepciones de justicia de los estudiantes acerca de cómo contribuye la evaluación justa en los métodos que se usan en la evaluación en el ámbito universitario. Para lograrlo, se llevó a cabo un doble estudio cuantitativo. El primero de los estudios, fue la aplicación de una escala adaptada de la *Selection Procedural Justice Scale* (Gilliland, 1993) a 385 estudiantes para determinar la actitud de los estudiantes hacia la justicia evaluativa. El segundo estudio que realizaron los profesores fue de corte experimental con dos grupos de estudiantes de la universidad de Southwestern. A través de ejercicios de simulación, se implementaron en los grupos las seis condiciones que pueden influir en una evaluación justa: explicitación de la evaluación alta o baja, proactividad alta o baja y retroalimentación alta o baja. Los resultados de ambos estudios muestran que la explicitación de los criterios de evaluación y calificación, una retroalimentación frecuente y la proactividad en el proceso evaluativo favorecen que los estudiantes conciban la evaluación como justa. Asimismo, desde la perspectiva de los estudiantes, la explicitación de los criterios así como la retroalimentación son los dos factores más influyentes en el desarrollo de una práctica evaluativa justa. Los resultados indican que existen tres aspectos claves para lograr una evaluación justa: la importancia de que los criterios de evaluación sean explícitos, la retroalimentación tiene que ser frecuente y la proactividad del docente en cuanto a la evaluación y la devolución de la información. Pepper y Pathak (2008) concluyen que es

fundamental que las evaluaciones sean imparciales y eso se refleje en las calificaciones para que estas sean más justas.

Otro estudio empírico posterior que arroja luz a esta cuestión que nos concierne es el estudio llevado a cabo por Tierney, Simon y Charland (2011) en Canadá. El propósito del estudio se centró en determinar cómo 77 docentes evaluaron a sus estudiantes en dos sistemas educativos canadienses diferentes con distintas políticas y estándares de evaluación. Relacionando la evaluación justa con un sistema de calificación basada en estándares (*Standards-Based Grading, SBG*), donde los estudiantes no son evaluados por igual sino en base a unos objetivos de competencias determinados, la investigación determina que: a) los docentes también tienen en cuenta el curso, el progreso de los estudiantes y los criterios para pasar a educación secundaria; b) las competencias en muchas ocasiones están relacionadas más con productos que con procedimientos; c) los docentes utilizan su criterio profesional para evaluar a sus estudiantes, no tan solo los grados basados en estándares; y d) los docentes consideran que sería justo evidenciar a los estudiantes porque les clasifican en los distintos grados. En síntesis, los autores concluyen que los docentes evalúan teniendo en cuenta lo que creen que es justo para sus estudiantes, para el bien individual y el común.

En concordancia con las investigaciones planteadas hasta el momento, la investigadora Nola Aitken (2012) de la Universidad de Lethbridge en Canadá se propone dar voz a los alumnos y alumnas. Para ello, realizó dos estudios de corte cualitativo realizados en 1994 y en 2007 con más de 250 estudiantes, y encontró que estos relacionan las prácticas de evaluación injustas con los siguientes aspectos de la práctica evaluativa:

- No existe relación entre el contenido y competencias enseñadas y aquello que se evalúa. La evaluación, no sirve para mejorar el aprendizaje porque tiene poca o escasa relación. Una prueba justa, desde la perspectiva de los alumnos, no tiene sorpresas.
- La práctica evaluativa justa se traduce en una variedad de oportunidades de evaluación para demostrar el conocimiento, la habilidad y la comprensión. Es necesario llevar a cabo diversas evaluaciones a lo largo del curso o el trimestre para que el resultado de las evaluaciones sea un claro reflejo de los conocimientos de los estudiantes (Stiggins, 2001).
- Las pruebas de evaluación deben ser múltiples y variadas, incluyendo distintos tipos de pruebas que supongan el uso de diferentes habilidades y conocimientos (Wiggins 1993).
- Las pruebas “cronometradas” o con un tiempo corto establecido en muchas ocasiones estresan a los estudiantes y no les dejan pensar con claridad. Por ello, es necesario flexibilizar el tiempo para realizar las pruebas de evaluación y que los estudiantes cuenten con un margen suficiente para expresar lo mejor de ellos.
- La implicación de los estudiantes en su propia evaluación es uno de los elementos claves para que ésta sea justa. Los alumnos quieren ser escuchados y que sus perspectivas se incluyan en la evaluación que van a realizar (Wiggins, 1993).

- La ponderación que se le asigna a cada prueba de evaluación es necesario que se adecue a la importancia que esa evaluación tenga en relación con la programación.
- Las instrucciones claras son fundamentales para ayudar a los estudiantes a demostrar sus conocimientos y habilidades durante la realización de la evaluación.
- Es necesario que las condiciones al realizar las evaluaciones sean favorables al estudiante: sin ruidos, sin molestias y en un ambiente adecuado para pensar.
- Las evaluaciones no deben servir para clasificar y etiquetar los estudiantes, ya que estos se sienten como ganadores o perdedores, como listos o tontos.
- La evaluación debe suponer una retroalimentación constante del proceso de aprendizaje de los estudiantes. La retroalimentación tiene que ser constructiva, siendo clara y oportuna para que los estudiantes puedan reflexionar sobre su éxito y en que pueden mejorar su trabajo.

El trabajo de Scott et al. (2014) es quizá el más reciente sobre la evaluación de estudiantes justa. Estos autores realizaron un estudio de enfoque mixto (cuantitativo-cualitativo) con el objetivo de profundizar en las concepciones de evaluación justa de los docentes. Los participantes fueron 3.312 docentes, estudiantes, familias y equipo directivo. Los resultados muestran que las prácticas evaluativas justas son más próximas al concepto de equidad y ofrecen cinco ideas clave sobre cómo debe ser la práctica evaluativa justa: a) Los educadores deben esforzarse por comprender el impacto que suponen sus

prácticas de evaluación a los estudiantes y sus familias; b) la evaluación debe ser diferenciada a cada estudiante, para atender a la capacidad, nivel socioeconómico, cultura y lengua de los estudiantes; c) todos los miembros de la comunidad educativa es necesario que denuncien prácticas indefendibles e ilógicas de evaluación; d) la frecuencia, la intensidad y la intrusión de las evaluaciones no debe abrumar ni a los estudiantes ni a sus familias; y e) finalmente, la evaluación no debe ser utilizada para contrarrestar el comportamiento inapropiado del estudiante o como recompensa deseada a un determinado comportamiento.

A modo de palabras de cierre, en cuanto al enfoque metodológico podemos afirmar que los estudios realizados para conocer la práctica o las concepciones de una evaluación justa tanto de docentes como de estudiantes son tanto cuantitativos y cualitativos y centrados en la educación obligatoria (tanto educación primaria como educación secundaria). En cuanto a los principales resultados, las investigaciones muestran que las concepciones de justicia en la evaluación están más cercanas a la equidad. Así, las prácticas evaluativas consideradas justas son aquellas que se adaptan a los estudiantes, comprenden su progreso y no utilizan únicamente las pruebas objetivas como instrumento para conocer qué han aprendido los alumnos y alumnas.

A propósito de los estudios realizados, Nola Aitken (2011) defiende la importancia de que en las investigaciones empíricas acerca de cómo es una evaluación justa se incluya la voz de los estudiantes. Para esta autora son los estudiantes los destinatarios últimos del proceso de enseñanza-aprendizaje y nadie mejor que ellos puede determinar cómo tiene que ser la evaluación. Aitken (2011) considera que además, no es

necesario únicamente escuchar su voz, sino confiar en ellos y darles más responsabilidades en el proceso evaluativo.

Por todo ello, es fundamental conocer cuáles son las concepciones de evaluación justa tanto de docentes como de alumnos y alumnas, de tal forma que la evaluación del aprendizaje de los estudiantes se convierta en un proceso más personal, individual y justo.



CAPÍTULO 5.

Metodología

En este capítulo presentamos la metodología utilizada para lograr los objetivos planteados en el apartado anterior. En primer lugar mostramos la metodología utilizada en la investigación previa, donde se realizó un estudio cuantitativo mediante el uso de Modelos Multinivel de tres de niveles. De este estudio presentamos el enfoque metodológico, la muestra del estudio, los instrumentos utilizados y el análisis de datos empleado.

En segundo lugar, describimos las características y peculiaridades del enfoque fenomenográfico para aproximarnos a las concepciones de docentes y estudiantes. A continuación, especificamos los participantes de nuestro estudio y presentamos la entrevista fenomenográfica como instrumento con el que hemos recogido la información. Por último, detallamos el proceso seguido para el análisis de los datos.

5.1. ESTUDIO PREVIO: INCIDENCIA DE LA EVALUACIÓN EN EL RENDIMIENTO

El estudio previo, como se ha señalado, aborda la incidencia de la evaluación en el rendimiento de los estudiantes mediante un análisis multinivel.

5.1.1. Enfoque metodológico

Tal como hemos presentado en el anterior apartado, la investigación previa realizada pretende determinar la relación entre la forma de evaluar por parte de los docentes de Educación Primaria en España y el desempeño académico de sus estudiantes en cuatro áreas: Competencia Lingüística, Competencia Matemática, Competencia Científica, y Competencia Social y Ciudadana.

Para alcanzar este objetivo se realizó una explotación especial de los datos de la Evaluación de Diagnóstico General (EDG) de Educación Primaria realizada en 2009 por el actual Instituto Nacional de Evaluación Educativa, del Ministerio de Educación de España (Instituto de Evaluación, 2010). Junto con las pruebas de rendimiento en cuatro competencias básicas, se aplicaron una serie de cuestionarios de contexto que aportaron información que ayuda no sólo a contextualizar los resultados de desempeño obtenidos, también permite profundizar en el conocimiento de las escuelas, los docentes y los estudiantes españoles de Educación Primaria.

5.1.2. Población, muestra y muestreo del estudio

Este estudio hace una explotación especial de la base de datos completa de Evaluación de Diagnóstico General de 2009 para Educación Primaria. Ello significa que utilizó datos de 1.341 docentes de dicha etapa que enseñan en 887 centros educativos,

y sus 28.708 estudiantes de 4º curso. El muestreo utilizado por el Instituto de Evaluación (actual INEE) para su obtención, en línea con otras evaluaciones internacionales, fue estratificado por conglomerados, con una estratificación constante de aproximadamente 50 centros en cada una de las Comunidades Autónomas.

5.1.3. Variables

Las variables utilizadas en el estudio fueron de tres tipos:

1. *Variables de producto.* Desempeño de los estudiantes en cuatro áreas: Competencia en Comunicación Lingüística, Competencia Matemática, Competencia en el Conocimiento y la interacción con el Mundo Físico, y Competencia Social y Ciudadana. Todas ellas estimadas mediante Teoría de Respuesta al Ítem (TRI) y con una puntuación en una escala de media 500 puntos y desviación típica 100.
2. *Variables explicativas.* Forma de evaluar por parte de los docentes, compuesto por ocho variables, todas ellas son variables de escala tipificadas:
 - a. Realizar una prueba sobre conocimientos previos antes de cada tema.
 - b. Realizar controles después de cada tema o lección.
 - c. Realizar controles escritos trimestrales.
 - d. Corregir los deberes y cuadernos.
 - e. Revisar las actividades que hacen en clase.
 - f. Realizar exámenes orales.

- g. Valorar la autoevaluación del alumnado.
 - h. Valorar el interés y la participación en clase de estudiantes.
3. *Variables de ajuste*. Características del estudiante: Índice socio-económico y cultural de la familia del estudiante (variable tipificada), Género del estudiante (hombre/mujer, variable *dummy*); y país de origen del estudiante (nativo/extranjero, variable *dummy*).

5.1.4. Instrumentos de recogida de información

Varios fueron los instrumentos que se aplican en la EDG y fueron utilizados por esta investigación. En primer lugar, cuatro pruebas de desempeño una para cada una de las cuatro competencias evaluadas: Competencia en Comunicación Lingüística (fiabilidad de 0,861), Competencia Matemática (0,800), Competencia en el Conocimiento y la interacción con el Mundo Físico (0,854) y Competencia Social y Ciudadana (0,856). En segundo lugar, un cuestionario dirigido a estudiantes y otro a familias, a partir del cual se elaboró el índice ISEC. Por último, se utilizó un cuestionario para docentes, del cual se obtuvieron los datos sobre su forma de evaluar.

5.1.5. Análisis de datos

Como estrategia de análisis de datos se utilizaron Modelos Multinivel de tres niveles (estudiante, aula y escuela), como alternativa metodológica más adecuada para abordar datos jerarquizados (Goldstein, 1995; Ma, Ma y Bradley, 2008; Martínez-Garrido y Murillo, 2013, 2014; Murillo, 2008). En palabras del profesor F. Javier Murillo (2008):

Los modelos multinivel constituyen la metodología de análisis más adecuada para tratar datos “jerarquizados” o “anidados” (por ejemplo, los estudiantes en aulas, o las aulas en escuelas), lo que la convierte en una estrategia imprescindible para la investigación educativa de carácter cuantitativo. Así, además de mejorar la calidad de los resultados, posibilita realizar análisis novedosos, tales como estimar la aportación de cada nivel de análisis (la del efecto del aula o la escuela) o las interacciones entre variables de distintos niveles. De esta forma se está en mejores condiciones de realizar estudios sobre factores asociados, sobre valor agregado o sobre equidad educativa. (p. 45)

Los modelos multinivel son ampliaciones de regresiones lineales, elaborando varios modelos de regresión en cada uno de los niveles de análisis de la investigación (Bickel, 2007; Murillo, 2008; Reise y Duan, 2003). En este sentido, los modelos de los distintos niveles están relacionados con variables explicativas del resto de niveles mediante los coeficientes de regresión.

Así, se realizaron cuatro de procesos de modelamiento multinivel con el software MLwiN 2.26, uno para cada una de las variables de producto, con los siguientes pasos:

1. *Modelo nulo*, sólo con la variable producto y el intercepto.
2. *Modelo ajustado*, con ISEC, género y país de origen.
3. *Modelo ajustado con las variables de evaluación* introducidas de forma individual.
4. *Modelo final*, con todas las variables incorporadas de forma simultánea, análogo al siguiente:

$$\begin{aligned}
 y_{ijk} &= b_{0jk} + b_{1jk} \text{ISEC}_{ijk} + b_{2jk} \text{Género}_{ijk} + b_{3jk} \text{Origen}_{ijk} + b_{4k} \text{Eval1}_{jk} + \dots \\
 &\dots + b_{10k} \text{Eval8}_{jk} + e_{ijk} \\
 \beta_{0jk} &= \beta_0 + v_{0k} + u_{0jk}
 \end{aligned}$$

$$\beta_{1jk} = \beta_1 + v_{1k} + u_{1jk} \dots \beta_{3jk} = \beta_3 + v_{3k} + u_{3jk}$$

$$\beta_{4k} = \beta_4 + u_{4k} \dots \beta_{10k} = \beta_{10} + u_{10k}$$

Con:

$$[e_{ijk}] \sim N(0, \Omega_e) : \Omega_e = [\sigma_{e_0}^2]$$

$$[u_{jk}] \sim N(0, \Omega_u) : \Omega_u = [\sigma_{u_0}^2]$$

$$[v_k] \sim N(0, \Omega_v) : \Omega_v = [\sigma_{v_0}^2]$$

Donde, para cada i= estudiante en cada j= aula, y k=escuela:

y_{ijk} es la cada una de las cuatro variables de producto:
Competencia Lingüística, Competencia Matemática,
Competencia Científica y Competencia Social y
Ciudadana.

$ISEC_{ijk}$ nivel socio-económico y cultural de la familia del
estudiante.

$Género_{ijk}$ género del estudiante (hombre/mujer).

$Origen_{ijk}$ país de origen del estudiante
(nativo/inmigrante).

$EvalI_{ij} \dots Eval8_{ij}$ cada una de las ocho variables de
evaluación del docente estudiadas.

e_{ijk} es el residual para cada estudiante.

u_{jk} el residual de cada aula.

v_k el residual de cada escuela.

5.2. INVESTIGACIÓN DE LAS CONCEPCIONES SOBRE UNA EVALUACIÓN JUSTA

Tras el estudio previo se desarrolló la investigación que conforma el núcleo fundamental de esta tesis centrada en las concepciones de evaluación justa y que se subdivide en tres subestudios:

1. Incidencia del contexto en las concepciones sobre evaluación.
2. Concepciones de docentes sobre una evaluación justa.
3. Concepciones de estudiantes sobre una práctica evaluativa justa (figura 5.1).

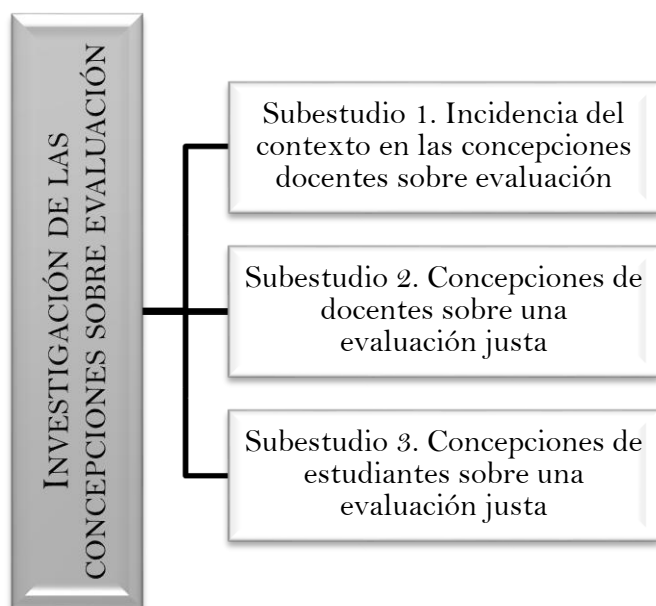


FIGURA 5.1. INVESTIGACIÓN DE LAS CONCEPCIONES. SUBESTUDIOS

Fuente: Elaboración propia.

5.2.1. Enfoque metodológico

Esta investigación utiliza la fenomenografía como metodología. Su selección está plenamente justificada el ser, posiblemente, la mejor forma de acercarse al conocimiento de las distintas concepciones sobre evaluación de docentes y estudiantes (Corbin y Strauss, 2008; Mertens, 2007; Taylor y Bogdan, 1996).

El enfoque fenomenográfico ha tenido una presencia discreta en investigaciones realizadas en el contexto español, quizás por su corto recorrido, ya que surgió hace menos de tres décadas o posiblemente por su ámbito tan focalizado de estudio en el proceso de adquisición de aprendizajes (Bowden y Walsh, 2000; González-Ugalde, 2014; Remesal, 2006; Richardson, 1999).

El método fenomenográfico se originó en la década de los ochenta en una investigación dirigida por Ference Marton en la Universidad de Göttemberg (Suecia). El profesor sueco, junto a su equipo de investigadores propuso una nueva metodología para aproximarse a las experiencias de aprendizaje y enseñanza en la formación universitaria. En palabras del propio Marton (1981) la fenomenografía se define como:

Un método de investigación para delinear las diferentes formas cualitativas en las cuales las personas experimentan, conceptualizan, perciben y comprenden varios aspectos de y en los fenómenos del mundo que los rodea. [...] se preocupa por las relaciones que existen entre los seres humanos y el mundo que los rodea. Se hace un esfuerzo por descubrir todas las comprensiones que la gente tiene sobre los fenómenos específicos y ordenarlas en categorías conceptuales. (p. 68)

Profundizando en la definición propuesta por Ference Marton (1981), la fenomenografía pretende conocer las formas, cualitativamente diferentes, en que las personas experimentan,

conceptualizan, perciben y comprenden el mundo que les rodea (Bowden, D'All'Alba, Martin, Masters, Laurillard, Marton, Ramsden y Stephanou, 1992; Orgill, 2012). Su propósito es descubrir las formas de comprensión que la gente tiene de fenómenos específicos para encuadrarlos dentro de categorías conceptuales. No realiza clasificaciones sobre el mundo tal cual, sino sobre las concepciones de las personas sobre el mundo. Así, la fenomenografía no pretende describir una realidad sino acercarse a cómo los sujetos experimentan un acontecimiento, proceso o hecho educativo (Booth, 1997; Marton, 1986; Reed, 2006).

Los supuestos teóricos básicos en los que se fundamenta la fenomenografía son: a) el conocimiento tiene un carácter relacional y holístico, b) las concepciones son la forma central del conocimiento, c) las descripciones son fundamentales para obtener un conocimiento científico acerca de concepciones, y d) el conocimiento científico centrado en las concepciones se basa en la diferenciación, la abstracción, la reducción y la comparación de los significados que los sujetos atribuyen a un determinado fenómeno (Booth, 1997; Marton, 1981, 1986).

El objeto central de la investigación fenomenográfica son las concepciones. El término “concepción” se utiliza a menudo para referirse a las “diferentes formas en que los individuos experimentan un elemento concreto de la realidad” (Sandberg, 1997, p.130). Las concepciones, de hecho, se consideran la unidad básica del conocimiento humano, y, si comprendemos las concepciones seremos capaces de profundizar en la psique de las distintas personas. Este supuesto ubica al conocimiento humano como una cuestión que se construye principalmente en el contexto social y cultural, fundamentándose en las

interacciones entre los sujetos u otros sujetos o el mundo (Reed, 2006).

Fijarse en las concepciones de los sujetos como aspecto central de un fenómeno y descripción del conocimiento significaba moverse desde una perspectiva objetivista e intersubjetiva de conocimiento hacia una visión más subjetivista, relativa y contextual de la construcción del mundo (Marton, 1996). En este sentido, existe un notable acuerdo en destacar la influencia de las concepciones en la conducta y acciones humanas y viceversa, en cómo el entorno y las interacciones determinan nuestras concepciones (Moyano-Díaz, 2016; Pozo, 2006; Remesal, 2006, 2009; Schraw y Olafson, 2003).

Siguiendo las ideas expuestas, podemos afirmar que el objetivo de la investigación fenomenográfica es una rica descripción y el objeto a describir son las concepciones de los sujetos participantes en la investigación. Marton (1986) matiza que realizar una fenomenografía supone identificar el número limitado de maneras cualitativamente diferentes de cómo diferentes personas vivencian, conceptualizan, perciben y entienden diversos tipos de fenómenos. Cuando se analizan las afirmaciones que las personas tienen acerca de un mismo fenómeno el resultado es un número limitado de maneras fundamentalmente diferentes de entender dicho fenómeno. Dentro de esta estructura, “el aprendizaje asume una importancia central, porque representa un cambio cualitativo entre una concepción interesada en algún aspecto particular de la realidad o bien en otro” (Marton, 1996, p. 53).

Margarita Mahncke (2010), en su tesis doctoral, profundizó en la existencia de diferentes tipos de Fenomenografía en función de la forma en que el investigador plantee la investigación y el

análisis de los datos resultantes. Para no excedernos en la exposición de los distintos tipos (hermenéutica, fenomenológica, discursiva o dialógica) consideramos que el presente estudio se enmarca en la Fenomenografía hermenéutica, la más cercana a la definición original de Johanson, Marton y Svensson (1985). Desde una perspectiva hermenéutica, la peculiaridad de este método de investigación reside en cómo el investigador interpreta las concepciones de los participantes del estudio, dotando a sus experiencias de un nuevo significado para lograr un abanico de distintas formas de experimentar un fenómeno educativo, especialmente aquellos relacionados con la enseñanza, el aprendizaje y la evaluación (Antón Nuño, 2012; Remesal, 2006, 2009).

En coherencia con el enfoque presentado, parece claro que la fenomenografía responde a la perfección al objetivo de conocer las concepciones sobre evaluación.

5.2.2. Participantes

Uno de los aspectos más relevantes del enfoque fenomenográfico es la selección de los participantes ya que es fundamental garantizar la mayor variación posible de una misma concepción. Centrarse en la variación de muestreo es una técnica que “permite seleccionar «casos críticos» que tienen que ser seleccionados a partir del juicio del investigador” (Cohen, Manion y Morrison, 2000, p. 103). La inclusión de estos casos críticos en la muestra seleccionada intencionalmente además, permitirá una mejor manifestación de las diversas formas de experimentar en este caso la evaluación justa. En la selección de los informantes con más probabilidades de proporcionar esta variación de formas de experimentar la evaluación es fundamental que el investigador

determine aquellos aspectos clave que pueden producir dicha variación (Trigwell, 2000; Walker, 1998).

En general, en la investigación cualitativa la selección de la muestra o participantes responde a un criterio de saturación de los datos (Glaser y Strauss, 1971). Dicha saturación teórica se produce cuando “no se encuentra datos adicionales a los ya hallados y el investigador confía en que una categoría está ya saturada empíricamente” (Glaser y Strauss, 1971, p. 65). Desde una perspectiva fenomenográfica, esta técnica no es la más adecuada, principalmente porque lo que buscamos son diferentes formas de concebir un fenómeno, independientemente de cuantas veces se repita esa misma concepción (Cohen et al., 2000).

Uniendo la necesidad de conseguir una variación adecuada de las distintas concepciones de evaluación justa y un número adecuado de sujetos se ha utilizado un muestreo no probabilístico por cuotas, fundamentado en la selección de aquellos individuos y criterios que garantizan mejor la variabilidad en los datos (McMillan y Schumacher, 2006; Patton, 1990).

Para ello, es necesario determinar los aspectos que pueden condicionar el fenómeno a estudiar y seleccionar a los participantes en base a ellos (las cuotas que ayudan a seleccionar los sujetos más adecuados para nuestra investigación). Como se ha indicado en los objetivos de la investigación, nos planteamos conocer las concepciones de evaluación en dos direcciones: en docentes y estudiantes. La elección de esta doble perspectiva de participantes responde a que tradicionalmente situamos a los profesores como informantes clave de los estudios cualitativos de evaluación

pero de forma menos frecuente a los estudiantes, los cuales son indudablemente los protagonistas del proceso de enseñanza-aprendizaje y los receptores finales de la evaluación (Black y Wiliam, 2009; Susinos, 2012).

Los participantes del estudio fenomenográfico son, en total 73 personas: 10 docentes en el primer subestudio, 31 docentes en el segundo y 32 estudiantes en el tercero. Fueron seleccionados mediante un muestreo no probabilístico por cuotas. Los criterios (o cuotas) utilizados para la selección tanto de docentes como de estudiantes son:

- *Nivel educativo que enseñan o cursan:* Educación Primaria y Educación Secundaria. Consideramos que las concepciones de evaluación pueden variar en función del nivel educativo, por la diferente formación de los y las docentes, y por la diferente edad y experiencia educativa en el caso de los y las estudiantes.
- *Titularidad del centro.* El segundo criterio de selección de los participantes es si los centros educativos son públicos o privados (tanto concertados como no concertados).
- *Contexto socio-económico del centro educativo.* La forma en que estudiantes y docentes experimentan y vivencian la evaluación puede cambiar en función del contexto en el que se encuentra su centro educativo. Por ello, se diferenció entre centros educativos ubicados en contextos favorables (nivel socioeconómico y cultural medio-alto) y otros en entornos desafiantes (nivel socioeconómico y cultural bajo). Para su selección se utilizaron tres criterios de forma complementaria: a) el

nivel socio cultural del barrio donde se encuentra el centro, b) el nivel socio-económico de las familias de los estudiantes que asisten a ese centro y c) las características sociales, culturales y lingüísticas de dichos estudiantes. Estos datos se han obtenido del Instituto Nacional de Estadística y de la Consejería de Educación de la Comunidad de Madrid.

El procedimiento de selección de los participantes fue el siguiente. En primer lugar se eligieron aquellos centros más adecuados en función de los tres criterios de selección anteriormente mencionados, de tal forma que tuviésemos centros escolares de distinta titularidad, nivel educativo y que estuvieran ubicados en diferentes contextos. Una vez elegidos los centros educativos, nos pusimos en contacto con el equipo directivo y les enviamos una carta con una explicación detallada de la investigación. Tras su aceptación, los equipos directivos de los centros nos ayudaron en la selección previa tanto de docentes como de estudiantes, con el criterio de que, en cuanto a los estudiantes, en educación Primaria fuesen de los últimos cursos y en educación secundaria del segundo ciclo de la Educación Secundaria Obligatoria. Los docentes podían ser de cualquier curso tanto en Educación Primaria como en Educación Secundaria.

Una vez realizada esta primera fase de selección, facilitamos una carta al equipo directivo para que se la enviasen a los docentes que habían pre-seleccionado y otra carta para ser enviada a las familias, donde se explicaba la investigación y se les solicitaba un acuerdo de consentimiento firmado para que sus hijos participasen en la misma, garantizando el anonimato y la fidelidad de los datos. Una vez el equipo directivo contó con

los acuerdos de consentimiento se eligieron entre 3 y 4 docentes y estudiantes de cada centro.

Como señalamos, en el primer subestudio participaron 10 docentes. En la tabla 5.1 se muestra la distribución en función de los tres criterios antes señalados y en la tabla 5.2 las características de estos diez profesores y profesoras participantes.

TABLA 5.1. PARTICIPANTES DEL SUBESTUDIO 1. DISTRIBUCIÓN DE LOS DOCENTES EN FUNCIÓN DEL CONTEXTO, DEL NIVEL EDUCATIVO Y DE LA TITULARIDAD DEL CENTRO

CONTEXTO	ED. PRIMARIA		ED. SECUNDARIA		TOTAL
	C. Públicos	C. Privados*	C. Públicos	C. Privados*	
Favorable	1	1	0	2	4
Desafiante	2	3	1	0	6
Total	3	4	1	2	10

Nota: * Centros privados, tanto concertados y como no concertados.

Fuente: Elaboración propia.

TABLA 5.2. CARACTERÍSTICAS DE LOS DOCENTES PARTICIPANTES EN EL SUBESTUDIO 1

DOCENTE	NIVEL EDUCATIVO	TITULARIDAD DEL CENTRO	CONTEXTO	ASIGNATURA	AÑOS EXP
D1	Primaria	Público	Desafiante	Tutor	10
D2	Primaria	Público	Desafiante	Matemáticas	14
D3	Primaria	Concertado	Desafiante	Música	3
D4	Primaria	Privado	Desafiante	Tutor	5
D5	Primaria	Concertado	Favorable	Tutor	8
D6	Primaria	Concertado	Desafiante	Tutor	12
D7	Secundaria	Privado	Favorable	Humanidades	10
D8	Secundaria	Concertado	Favorable	Lengua	15
D9	Secundaria	Público	Desafiante	Historia	22
D10	Secundaria	Privado	Favorable	Lengua	8

Fuente: Elaboración propia.

En el subestudio 2 participaron un total de 31 docentes, 15 de Educación Primaria y 16 de Educación Secundaria. En la tabla 5.3 aparece la distribución por cuotas.

TABLA 5.3. PARTICIPANTES DEL SUBESTUDIO 2. DISTRIBUCIÓN DE LOS DOCENTES EN FUNCIÓN DEL CONTEXTO, DEL NIVEL EDUCATIVO Y DE LA TITULARIDAD DEL CENTRO

CONTEXTO	ED. PRIMARIA		ED. SECUNDARIA		TOTAL
	C. Públicos	C. Privados*	C. Públicos	C. Privados*	
Favorable	3	5	4	4	16
Desafiante	4	3	4	2	15
Total	7	8	8	6	31

Nota: * Centros privados, tanto concertados y como no concertados.

Fuente: Elaboración propia.

En la tabla 5.4 se describen algunas de las características de estos docentes tales como su especialidad o sus años de experiencia docente.

TABLA 5.4. CARACTERÍSTICAS DE LOS DOCENTES PARTICIPANTES EN EL SUBESTUDIO 2

DOCENTE	NIVEL EDUCATIVO	TITULARIDAD DEL CENTRO	CONTEXTO	ASIGNATURA	AÑOS EXP
D1	Primaria	Público	Desafiante	Tutor	10
D2	Primaria	Público	Desafiante	Matemáticas	14
D3	Primaria	Privado	Desafiante	Tutor	3
D4	Primaria	Privado	Desafiante	Tutor	5
D6	Primaria	Privado	Desafiante	Tutor	12
D9	Secundaria	Público	Desafiante	Historia	25
D11	Secundaria	Privado	Desafiante	Inglés	4
D16	Secundaria	Público	Desafiante	Matemáticas	5
D17	Secundaria	Privado	Desafiante	Tecnología	3
D18	Secundaria	Público	Desafiante	Inglés	7
D19	Primaria	Privado	Desafiante	Lengua	3
D21	Secundaria	Privado	Desafiante	Biología	4
D28	Primaria	Público	Desafiante	Tutor	9
D29	Secundaria	Público	Desafiante	Plástica	15
D31	Secundaria	Privado	Desafiante	Ciencias	15
D5	Primaria	Público	Favorable	Tutor	8
D7	Secundaria	Privado	Favorable	Sociales	10
D8	Secundaria	Privado	Favorable	Lengua	25
D10	Secundaria	Privado	Favorable	Francés	8
D12	Primaria	Público	Favorable	Lengua	5
D13	Secundaria	Público	Favorable	Dibujo	20
D14	Secundaria	Público	Favorable	Sociales	4
D15	Secundaria	Público	Favorable	Biología	25
D20	Secundaria	Público	Favorable	Lengua	19
D22	Primaria	Público	Favorable	Tutor	7
D23	Primaria	Público	Favorable	PT	15
D24	Secundaria	Privado	Favorable	Música	21
D25	Primaria	Privado	Favorable	Tutora	39
D26	Secundaria	Privado	Favorable	Tecnología	20
D27	Primaria	Privado	Favorable	Tutor	25
D30	Primaria	Privado	Favorable	Tutor/PT	20

Nota: * Centros privados, tanto concertados y como no concertados.

Fuente: Elaboración propia.

En el subestudio 3, por último, participaron 32 estudiantes, 15 de 6º curso de Educación Primaria y 17 de 4º curso de Educación Secundaria Obligatoria. En la tabla 5.5 se muestra

su distribución en función del nivel educativo, titularidad del centro y contexto en el que se ubica su escuela.

En la tabla 5.6 se detalla las características de los estudiantes participantes en el estudio.

TABLA 5.5. PARTICIPANTES DEL SUBESTUDIO 3. DISTRIBUCIÓN DE LOS ESTUDIANTES EN FUNCIÓN DEL CONTEXTO, DEL NIVEL EDUCATIVO Y DE LA TITULARIDAD DEL CENTRO

CONTEXTO	ED. PRIMARIA		ED. SECUNDARIA		TOTAL
	C. Públicos	C. Privados*	C. Públicos	C. Privados*	
Favorable	3	3	4	4	14
Desafiante	5	4	5	4	18
Total	8	7	9	8	32

Nota: * Centros privados, tanto concertados y como no concertados.

Fuente: Elaboración propia.

TABLA 5.6. CARACTERÍSTICAS DE LOS DOCENTES PARTICIPANTES EN EL SUBESTUDIO 3

ESTUDIANTE	NIVEL EDUCATIVO	TITULARIDAD DEL CENTRO	CONTEXTO	EDAD
E1	Primaria	Público	Desafiante	11
E4	Primaria	Privado	Desafiante	11
E5	Primaria	Privado	Desafiante	12
E6	Primaria	Privado	Desafiante	12
E9	Secundaria	Público	Desafiante	15
E20	Secundaria	Público	Desafiante	15
E21	Secundaria	Público	Desafiante	16
E22	Secundaria	Público	Desafiante	16
E23	Secundaria	Público	Desafiante	17
E24	Primaria	Privado	Desafiante	11
E25	Primaria	Público	Desafiante	12
E26	Primaria	Público	Desafiante	12
E27	Primaria	Público	Desafiante	13
E28	Primaria	Público	Desafiante	12
E29	Secundaria	Privado	Desafiante	16
E30	Secundaria	Privado	Desafiante	15
E31	Secundaria	Privado	Desafiante	17
E32	Secundaria	Privado	Desafiante	16
E2	Primaria	Privado	Favorable	12
E3	Secundaria	Público	Favorable	16
E7	Secundaria	Privado	Favorable	16
E8	Secundaria	Privado	Favorable	15
E10	Secundaria	Público	Favorable	16
E11	Secundaria	Público	Favorable	16
E12	Secundaria	Público	Favorable	15
E13	Secundaria	Privado	Favorable	15
E14	Secundaria	Privado	Favorable	16
E15	Primaria	Público	Favorable	11
E16	Primaria	Público	Favorable	12
E17	Primaria	Público	Favorable	12
E18	Primaria	Privado	Favorable	11

Fuente: Elaboración propia.

El trabajo de campo fue desarrollado entre los meses de septiembre de 2015 y junio de 2016 con el apoyo y ayuda del equipo directivo así como la implicación de todos los sujetos participantes en la investigación.

5.2.3. Instrumento de recogida de información: la entrevista fenomenográfica

La técnica de recogida de información utilizada en la investigación ha sido la entrevista fenomenográfica basada en el auto-relato (Trigwell, 1997, 2000), técnica más utilizada para la recopilación de información fenomenográfica. Desde la perspectiva de Marton (1996) la entrevista fenomenográfica es una producción interactiva donde, “las experiencias y concepciones se construyen de forma conjunta entre el entrevistador y el entrevistado” (p. 4427). Para ello, es fundamental que la entrevista tenga una o pocas preguntas predeterminadas y que el investigador tenga siempre presente el propósito del estudio para que guíe la entrevista (Bowden, 1996; Dortins, 2002; Marton, 1986). En este caso, el investigador ha profundizado a través de las entrevistas en las concepciones de los estudiantes acerca de qué es una evaluación justa y sus implicaciones. Asimismo, el investigador no debe posicionarse frente a un determinado tipo de concepción, sino ser neutral y mostrar interés compartido con los participantes del estudio.

Las experiencias captadas por la entrevista se constituyen conjuntamente entre el entrevistador y el entrevistado (Marton, 1996) y se tematiza a través de una “conversación entre dos sujetos sobre un tema de interés mutuo” (Kvale, 1996, p. 125). Säljö (1996, p. 23) matiza que “este tema de interés compartido es el tema que guía todo el discurso”.

La entrevista típica fenomenográfica es de carácter semi-estructurado con solo unas pocas preguntas clave predeterminada. Este aspecto es fundamental para no dirigir las concepciones de los sujetos entrevistados. Este aspecto contrasta con la entrevista cualitativa típica donde existe un

guión pre-determinado que la pauta y que está conformando por distintas preguntas. En la entrevista fenomenográfica el objeto de estudio (la evaluación justa) enfoca al entrevistador en todo momento y guía la situación de la entrevista. Säljö (1996, 1997) señala que los investigadores tienen que tener cuidado en no tener prejuicios frente a lo que narra el sujeto entrevistado para no sesgar su discurso, motivo por el cual es importante que las intervenciones del investigador guíen las ideas expresadas por el participante entrevistado. En palabras de Remesal (2006, p. 107) “entrevistador y entrevistado interaccionan desde el primer momento hasta que se despiden. Se influyen mutuamente y en cierto sentido, son coautores de la entrevista resultante”.

En el primer estudio, las preguntas que guiaron la investigación fueron: *¿Cómo evalúas?* y *¿Cómo sería tu evaluación ideal?* *¿Concibes esta evaluación como justa?*

En los subestudios 2 y 3, siguiendo las premisas teóricas, las entrevistas solo tuvieron una pregunta generadora: *¿Qué es para ti una evaluación justa?*

Las entrevistas realizadas tuvieron una duración de entre 15 y 20 minutos y se realizaron en un contexto y espacio donde el entrevistado se sintiese cómodo. En todos los casos, las entrevistas se realizaron en los centros educativos de los participantes y se grabaron por completo. De todas las entrevistas realizadas solo encontramos la negativa de ser grabada en una de las docentes, la cual solicitó que sus ideas fuesen escritas por el investigador pero no registradas por la grabadora.

La presente tesis doctoral se ajusta a las normas de investigación acordadas por el Comité de Ética de la

Investigación de la Universidad Autónoma de Madrid con el fin de proteger los derechos fundamentales de las personas y los compromisos morales por los Estatutos de la Universidad. Toda la información obtenida en esta tesis ha sido tratada de forma confidencial y conforme a la Ley sobre Protección de Datos Personales (Ley 15/1999).

En este sentido, todos los participantes fueron informados de los objetivos de la investigación debidamente, y se solicitó su consentimiento para ser grabados. En el caso de estudiantes menores de edad, se solicitó a las familias un acuerdo de consentimiento firmado para que sus hijos participasen en la misma. Tanto para los docentes como para los estudiantes se mantuvo escrupulosamente su anonimato y fuimos extremadamente cautelosos de no alterar sus palabras ni sacarlas del contexto.

5.2.4. Análisis de datos y categorías de análisis del estudio

De acuerdo con el método fenomenográfico, se ha utilizado un criterio de proximidad al discurso de los participantes propuesto por Marton (1986). Como podemos apreciar en la figura 5.2, el análisis de los datos ha constado de siete fases para examinar las concepciones de los estudiantes y los docentes. A continuación se detallan cada una ellas (figura 5.2).

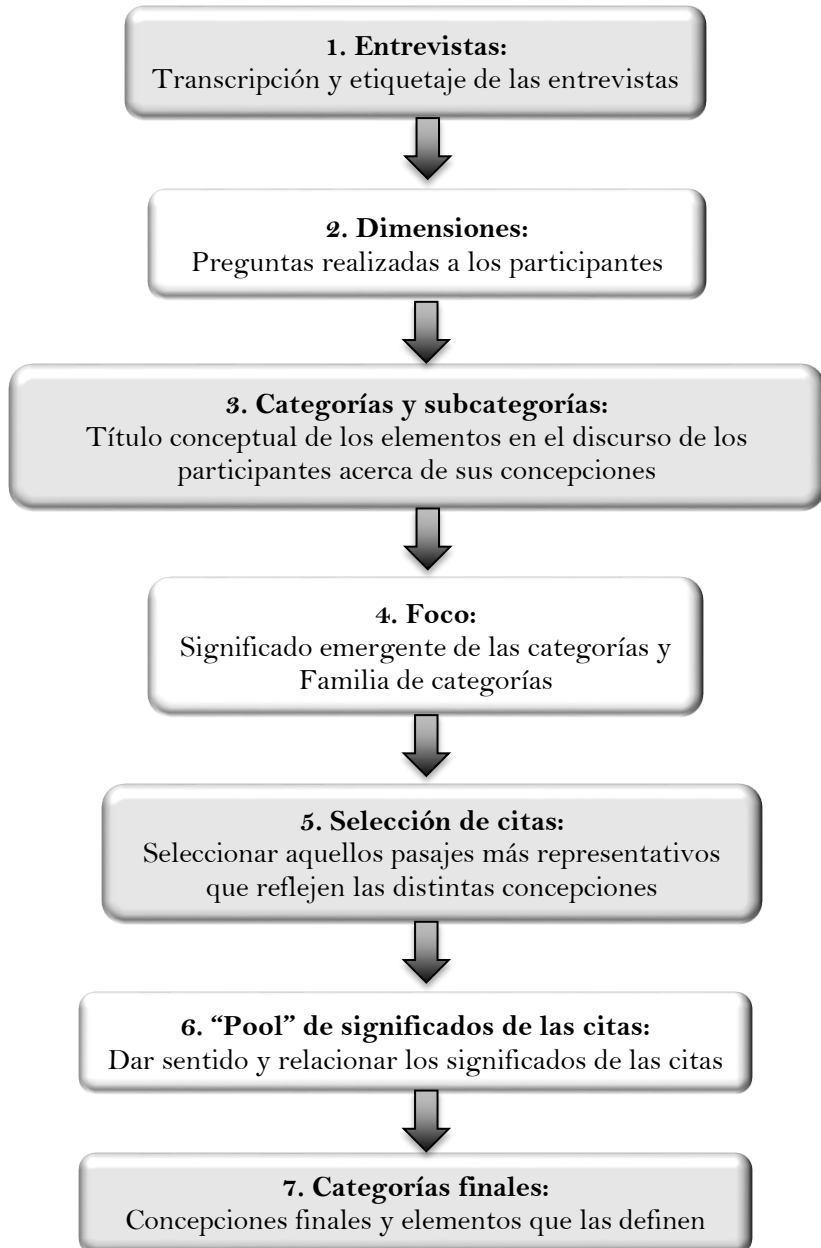


FIGURA 5.2. PROCESO FENOMENOGRÁFICO SEGUIDO EN EL ANÁLISIS DE LOS DATOS

Fuente: Elaboración propia.

Fase 1. Una vez realizadas las 63 entrevistas, se transcribieron de forma completa con la ayuda de software F4, poniendo especial atención en ser fieles a las palabras de los participantes. Una vez transcritas les asignamos una etiqueta a cada una de ellas con el siguiente formato [E1: Pri, Pub, Des, p.1]. En él aparecen cuatro informaciones: a) el número de estudiante o docente, b) el nivel educativo (Pri: educación primaria, y Sec: educación secundaria); c) titularidad del centro (Pu: público y Pr: privado), d) contexto donde se ubica el centro (Fav: favorable y Des: desafiante), y e) el número de la página donde se encuentra la cita. Las etiquetas asociadas a cada entrevista permiten, por un lado, mantener el anonimato de la entrevista, y, por otro, tener presentes los criterios que nos hemos planteado para garantizar la variación de las concepciones.

Fase 2. Con las entrevistas transcritas en su totalidad se realizó una lectura exhaustiva y minuciosa para identificar aquellas palabras y expresiones más utilizadas por los participantes referentes a cómo conciben una evaluación justa (Marton, 1986; Sandbergh, 1997). Estas palabras o expresiones seleccionadas se iban trasladando a otro archivo para tener siempre presentes aquellos aspectos destacados de las entrevistas, permitiéndonos tener una visión general de los resultados que íbamos obteniendo en todo momento. En este primer acercamiento del investigador a los datos es fundamental tener en cuenta el horquillado en el análisis de las entrevistas (Marton, 1986). En otras palabras, el horquillado significa que el investigador debe acercarse tanto a la entrevista y los datos que se analizaron con la mente abierta y sin participación alguna de sus perspectivas. Este procedimiento es necesario que se haga con cuidado y rigor estando libre de suposiciones y de ideas previas que puedan inducir al sesgo o al error al analizar las entrevistas.

Fase 3. A continuación, se conjugaron y organizaron las expresiones seleccionadas para crear las categorías iniciales o pre-categorías de su discurso. Las categorías creadas surgieron de las propias palabras de los docentes, favoreciendo que el análisis sea lo más fiel posible al discurso y a las concepciones que presentaban los sujetos entrevistados. Posteriormente, se compararon las concepciones de un docente con el resto, viendo regularidades y creando pre-categorías cada vez más amplias, analizando, asimismo, la relación entre las distintas pre-categorías (Marton, 1981, 1986). De acuerdo con Sjostrom y Dahlgren (2002), basamos el análisis de las pre-categorías de acuerdo a tres principios:

- *Frecuencia*, relativo a cuantas veces los docentes mencionan esa idea o concepción relacionada con una evaluación justa.
- *Posición*, donde se ubica la concepción respecto al resto de concepciones, identificando aquellas más relevantes.
- *Pregnancy*, analizando aquellos aspectos en los que los docentes entrevistados ponen más énfasis o consideran más importante para desarrollar una evaluación justa.

En el análisis de las entrevistas fenomenográficas es fundamental ser consciente de que existe un número limitado de formas de concebir la evaluación justa (Trigwell, 2000). Dicho de otro modo, para cualquier fenómeno, hay un número limitado de formas cualitativamente diferentes de experimentarlo y expresarlo en forma de concepción. Por ello es importante no sentirse presionado por encontrar más variaciones en las concepciones, sino más bien por explicar de la forma más detallada y concreta posible cada variación. De hecho, no es tan importante el número de categorías en la

fenomenografía sino más bien la diferencias y similitudes que sirven para vincular y diferenciar una concepción de otra, es decir, la estructura y significado relacionado entre las distintas categorías (Reed, 2006).

Fase 4. En cuarto lugar, se organizaron las pre-categorías en familias de tal forma que las concepciones queden organizadas en un árbol de significado coherente con las concepciones de los informantes.

Fase 5. Se seleccionaron las citas más representativas de cada una de las pre-categorías establecidas en la fase 2, poniendo especial atención en ser fieles a las palabras y el sentido de las concepciones de los estudiantes.

Fase 6. Se comparó cada una de las citas seleccionadas con las otras citas, creando un *pool of meaning* de todas las citas de los estudiantes. Las citas seleccionadas del discurso de los participantes conforman el *pool* de datos que forma la base para el siguiente paso importante en el análisis. La atención del investigador avanza de los sujetos individuales (es decir, desde las entrevistas a partir de los cuales se extrajeron los fragmentos más relevantes) al significado de cada una de estas citas (Harris, 2008). A continuación se elaboró el *pool* de significado, permitiendo de esta manera dar un sentido concreto a las citas y dotándoles de un significado real al compararlas con otros fragmentos relevantes acerca de cómo tiene que ser una práctica evaluativa justa para los participantes (Harris, 2008).

Fase 7. Por último, las pre-categorías se analizaron en los distintos *pools* de significado que surgieron del análisis. Estos *pools* de significado fueron estudiados de forma individual y después de forma cruzada, estableciendo las categorías finales y representando las distintas formas de concebir la evaluación

justa. Cada una de las categorías referentes a las distintas concepciones es ilustrada por distintas citas, lo que nos aproxima a una comprensión detallada y llena de matices de cada una de las concepciones (Harris, 2008; Marton, 1986; Reed, 2006).

El análisis de los datos se realizó mediante al apoyo del software Atlas.ti (versión 7) para analizar y clasificar los datos de forma clara y precisa. El programa ayuda a organizar las entrevistas a examinar, creando dos unidades hermenéuticas (una para las entrevistas de los profesores y profesoras y otra para las realizadas con el alumnado). En cada una de las unidades hermenéutica, el Atlas.ti ayudó a categorizar (o codificar) las entrevistas así como jerarquizar los datos (esenciales para elaborar los *pool* de significados de los que hemos estado hablando). Asimismo, el software permitió crear redes semánticas donde se muestran las relaciones entre las categorías y subcategorías generadas del análisis de las distintas concepciones de docentes y estudiantes. Estos gráficos, que facilitan la visualización e interpretación de los datos, se incluyen en los resultados de la presente tesis doctoral.



CAPÍTULO 6.

Resultados

En este capítulo se presentan los resultados obtenidos en esta investigación. En primer lugar, se muestran los hallazgos del estudio previo acerca de la incidencia de la práctica evaluativa en el aprendizaje de los estudiantes. A continuación se exponen los resultados de la investigación de las concepciones sobre evaluación. Para organizar los resultados de dicha investigación, se presentan de forma separada los hallazgos de cada uno de los tres subestudios centrados en conocer la incidencia del contexto en las concepciones de evaluación y las concepciones del profesorado y los estudiantes sobre una evaluación justa.

6.1. ESTUDIO PREVIO: INCIDENCIA DE LA EVALUACIÓN EN EL RENDIMIENTO

Como se ha comentado de forma reiterada a lo largo de este volumen, antes de realizar la investigación sobre concepciones propiamente dicha se realizó un estudio previo cuyo objetivo es determinar la incidencia de la evaluación en el rendimiento de los estudiantes. Para ello se llevó a cabo un Análisis Multinivel siendo diferentes medidas del desempeño académico las variables dependientes, y distintas prácticas evaluativas las variables explicativas.

El rendimiento de los estudiantes se midió a partir de su desempeño en Competencia Lingüística, Competencia Matemática, Competencia Científica, y Competencia Social y Ciudadana.

Las variables explicativas sobre evaluación con las que trabajamos son:

- Realizar una prueba sobre conocimientos previos antes de cada tema
- Realizar controles después de cada tema o lección
- Realizar controles escritos trimestrales
- Corregir los deberes y cuadernos
- Revisar las actividades que hacen en clase
- Realizar exámenes orales
- Valorar la autoevaluación del alumnado
- Valorar el interés y la participación en clase de estudiantes

Para poder alcanzar nuestro objetivo, elaboramos los siguientes cuatro modelos tal como hemos planteado en la metodología:

1. *Modelo nulo*, donde sólo añadimos la variable producto y el intercepto.
2. *Modelo ajustado*, con las variables ISEC, género y país de origen.
3. *Modelo ajustado con las variables de evaluación* introducidas de forma individual.
4. *Modelo final*, con todas las variables incorporadas de forma simultánea, análogo al siguiente:

Una vez elaborados los modelos multinivel nulo y ajustado, se introdujeron al modelo individualmente las ocho variables explicativas en cada uno de los cuatro modelos ajustados que se han mostrado en la tabla 6.1. Con ellos se determinan cuáles son las actividades de evaluación que impactan de manera significativa en el desarrollo de los estudiantes.

Del análisis de la tabla 6.1 se extraen una serie de interesantes ideas para cada una de las áreas evaluadas:

- *Competencia Lingüística*. De las ocho variables explicativas estudiadas relativas a la evaluación, sólo una parece tener efecto en el desempeño en competencia lingüística -descontando el efecto del ISEC y el origen-, el hecho de que el docente corrija los deberes y cuadernos. Concretamente, los datos indican que por cada punto que aumente (o disminuya) sobre la media la variable Corregir los deberes y cuadernos, el desarrollo

de la Competencia Lingüística de los estudiantes aumenta (o disminuye) en 3,3 puntos.

TABLA 6.1. RESULTADOS DE LOS MODELOS MULTINIVEL CONSTRUIDOS A PARTIR DE LA INTRODUCCIÓN DE CADA VARIABLE EXPLICATIVA EN LOS MODELOS AJUSTADOS

	COMPETENCIA LINGÜÍSTICA	COMPETENCIA MATEMÁTICA	COMPETENCIA CIENTÍFICA	COMPETENCIA SOCIAL Y CIUDADANA
	B (EE)	B (EE)	B (EE)	B (EE)
Realizar una prueba sobre conocimientos previos antes de cada tema	NS	NS	NS	NS
Realizar controles después de cada tema o lección	NS	NS	4,20 (1,61)	4,67 (1,57)
Realizar controles escritos trimestrales	NS	-1,72 (0,80)	NS	-1,68 (0,82)
Corregir los deberes y cuadernos	3,32 (1,16)	NS	4,58 (1,74)	3,92 (1,72)
Revisar las actividades que hacen en clase	NS	NS	4,87 (2,00)	NS
Realizar exámenes orales	NS	NS	NS	NS
Valorar la autoevaluación del alumnado	NS	NS	NS	NS
Valorar el interés y la participación en clase de estudiantes	NS	NS	NS	4,29 (2,08)

Nota: NS Aportación no significativa con $\alpha=0,05$.

Fuente: Elaboración propia.

- *Competencia Matemática.* De nuevo, una sola variable explicativa parece realizar una aportación

estadísticamente significativa al modelo ajustado para esta variable producto: Realizar controles trimestrales. Y, curiosamente, el índice es negativo, lo que significa que el hecho de que el docente realice controles trimestrales tiene una incidencia negativa en el desarrollo de la Competencia Matemática. Expresado en números, por cada punto que el docente aumente (o disminuya) la frecuencia de la realización de controles trimestrales, el desempeño en la variable producto Competencia Matemática de los estudiantes disminuye (o aumenta) 1,72 puntos.

- *Competencia Científica.* En tres de los ocho modelos multinivel calculados con esta variable de producto, uno para cada variable explicativa introducida individualmente, la variable de evaluación hace una aportación estadísticamente significativa, en concreto, con las variables: Corregir los deberes y cuadernos, Realizar controles después de cada tema, y Revisar las actividades realizadas en clase. Ello implica que estas tres variables inciden en el desempeño en la Competencia Científica. Específicamente, los datos muestran que por cada punto que aumente (o disminuya) la frecuencia con la que el docente revisa las actividades realizadas en clase, el estudiante aumenta (o disminuye) en 4,87 puntos su desarrollo en esta competencia, 4,58 por la frecuencia en corregir los deberes y cuadernos, y 4,2 puntos por la frecuencia en realizar controles después de cada tema.

- *Competencia Social y Ciudadana*. Son cuatro las variables explicativas que, según los datos, influyen de manera significativa en el desarrollo de esta competencia: Realizar controles después de cada tema o lección, Realizar controles escritos trimestrales, Corregir los deberes y cuadernos, y Valorar el interés y la participación en clase de estudiantes. Todos ellos con una relación positiva; excepto realizar controles escritos trimestrales que, al igual que ocurría con el desempeño en Competencia Matemática, tiene una relación negativa. En los tres casos positivos la incidencia es de más 4 puntos, y en el caso del coeficiente negativo de 1,7.

Una vez discriminado las variables que realizan una aportación significativa individualmente a los modelos ajustados, para cada variable de producto se estima un modelo con todas ellas de forma simultánea (el modelo final) con el fin de obtener un mayor estimador de máxima verosimilitud en el modelo. El estudio de la variación de la razón de verosimilitud refleja que prácticamente no varía para ninguna de las variables. Así pues, el último paso es el cálculo de los cuatro modelos multinivel final, con las variables de ajuste y todas las variables explicativas de forma simultánea (tabla 6.2). Para las variables producto Competencia Lingüística y Competencia Matemática, al tener solo una variable explicativa con aportación significativa en los modelos ajustados, estos coinciden con los modelos finales. Luego para estas dos variables estos modelos finales no aportarían más información. Para las otras dos variables de producto, sin embargo, los modelos finales aportan

información sobre las variables que inciden más en el desarrollo de las competencias, y las que hacen una aportación diferente.

TABLA 6.2. RESULTADOS DE LOS MODELOS MULTINIVEL FINALES

	COMPETENCIA LINGÜÍSTICA	COMPETENCIA MATEMÁTICA	COMPETENCIA CIENTÍFICA	COMPETENCIA SOCIAL Y CIUDADANA
	B (EE)	B (EE)	B (EE)	B (EE)
Parte fija				
Intercepto	492,33 (4,73)	503,57 (1,45)	481,13 (1,60)	479,34 (6,96)
ISEC	28,64 (0,69)	26,44 (0,65)	26,63 (0,63)	25,12 (0,68)
Origen (nativo- inmigrante)	-4,10 (2,01)	-5,04 (2,10)	-4,75 (2,03)	NS
Variables de proceso				
Realizar controles después de cada tema o lección			3,73 (1,62)	4,32 (1,16)
Corregir los deberes y cuadernos	3,32 (1,16)		4,26 (2,00)	
Realizar controles escritos trimestrales		-1,67 (0,80)		-1,18 (0,82)
Valorar el interés y la participación en clase de estudiantes				4,34 (2,10)
Parte aleatoria				
Entre escuelas	774,88 (57,18)	785,66 (58,46)	1277,23 (76,56)	987,30 (66,48)
Entre aulas	33,35 (28,93)	35,84 (29,94)	0,00 (00,00)	5,94 (27,16)
Entre estudiantes	7856,91 (72,42)	8104,61 (74,70)	7488,80 (68,46)	7878,70 (72,80)

Nota: NS Aportación no significativa con $\alpha=0,05$.

Fuente: Elaboración propia.

Analizando la tabla 6.2 se observa que, para Competencia Científica, de las tres variables que tuvieron una aportación significativa, individualmente se "cae" la variable Revisar las actividades que realizan en casa. Para la variable producto

Competencia Social y Ciudadana de las cuatro variables que realizan aportaciones en los modelos individuales, y tomadas conjuntamente solo tres de ellas realiza una aportación significativa; se pierde, por tanto, la variable corregir los deberes.

Con todo ello, se puede concluir que las actividades de evaluación que realiza el docente en el aula tiene una incidencia en el desarrollo global del estudiante controlando el efecto del nivel socio-económico y el origen (tabla 6.3). Más exactamente, tres son las actividades evaluativas que han mostrado su relación positiva con el desempeño en alguna o en varias de las competencias estudiadas:

- Realizar controles después de cada tema o lección, que incide sobre el desarrollo de su Competencia Científica y en el desarrollo de su Competencia Social y Ciudadana.
- Corregir los deberes y cuadernos, que impacta sobre el desempeño de los estudiantes en Competencia Lingüística y en Competencia Científica.
- Valorar el interés y la participación de los estudiantes en clase, que afecta de manera positiva sobre el desarrollo de su Competencia Social y Ciudadana.

Por el contrario, llevar a cabo controles escritos trimestrales afecta negativamente en el desarrollo de los estudiantes en Competencia Matemática y en Competencia Social y Ciudadana.

Por último, hay tres actividades de evaluación que según estos datos no parecen estar directamente asociadas con el desarrollo de ninguna competencia:

- Realizar una prueba sobre conocimientos previos antes de cada tema.
- Realizar exámenes orales.
- Valorar la autoevaluación del alumnado.

TABLA 6.3. RESUMEN DEL PROCESO DE MODELADO MULTINIVEL

	C. LINGÜÍSTICA	C. MATEMÁTICA	C. CIENTÍFICA	C. SOCIAL Y CIUDADANA
Realizar una prueba sobre conocimientos previos antes de cada tema				
Realizar controles después de cada tema o lección			XX	XX
Realizar controles escritos trimestrales		XX*		XX*
Corregir los deberes y cuadernos	XX		XX	X
Revisar las actividades que hacen en clase			X	
Realizar exámenes orales				
Valorar la autoevaluación del alumnado				
Valorar el interés y la participación en clase de estudiantes				XX

Notas: X variables con una aportación significativa en los modelos con variables introducidas individualmente

XX variables con una aportación significativa en los modelos finales.

(*) aportación negativa.

Fuente: Elaboración propia.

Vista la incidencia de algunas actividades de evaluación sobre el desarrollo de las diferentes competencias, puede ser interesante analizar qué es lo que hacen los docentes de Educación

Primaria en España en las otras variables sobre diferentes actividades de evaluación. De esta forma será posible valorar en qué medida las prácticas evaluativas que se han mostrado especialmente eficaces son desarrolladas por los maestros y las maestras en España. Los datos globales aparecen en la tabla 6.4.

Los datos aportados por los 1.341 docentes estudiados muestran con claridad las actividades que desarrollan respecto a la evaluación de sus estudiantes. En una rápida mirada es posible agrupar las prácticas en tres grandes grupos en función de su frecuencia de uso.

TABLA 6.4. DISTRIBUCIÓN DE LOS DOCENTES SEGÚN LA FRECUENCIA DE REALIZACIÓN DE DIFERENTES ACTIVIDADES DE EVALUACIÓN. DATOS EN PORCENTAJE

	NUNCA O CASI NUNCA	ALGUNA S VECES	CASI SIEMPRE	SIEMPRE
Realizar una prueba sobre conocimientos previos antes de cada tema	31,15	37,78	22,01	9,06
Realizar controles después de cada tema o lección*	,90	4,51	23,67	70,92
Realizar controles escritos trimestrales*	45,57	22,52	11,96	19,95
Corregir los deberes y cuadernos*	,08	2,18	23,98	73,76
Revisar las actividades que hacen en clase	,08	1,13	24,00	74,79
Realizar exámenes orales	12,80	62,58	17,92	6,70
Valorar la autoevaluación del alumnado	12,76	42,14	28,40	16,70
Valorar el interés y la participación en clase de alumnas/os*	,00	1,20	16,69	82,11

Nota: (*) variables asociadas con el desempeño.

Fuente: Elaboración propia a partir de los datos de le EGD 2009.

Por una parte están las prácticas de evaluación desarrolladas por la mayoría de los docentes de Primaria, que serían cuatro:

Realizar controles después de cada tema o lección, Corregir los deberes y cuadernos, Revisar las actividades hechas en clase, y Valorar para la calificación el interés y la participación en clase de los estudiantes. En todos ellas, más del 70% de los maestros y maestras afirman hacerlo siempre (el 70, 9% Realizar controles después de cada tema o lección, el 73,7% Corregir los deberes y cuadernos, y el 74,8% valorar para la calificación el interés y la participación en clase de los estudiantes), y más del 96% de ellos dicen hacerlo siempre o casi siempre (Realizar controles después de cada tema o lección el 96,6%, Corregir los deberes y cuadernos el 97,7%, Revisar las actividades hechas en clase el 98,8%, y Valorar el interés y la participación en clase de los estudiantes el 98,8%).

En el segundo grupo estarían las actividades de evaluación realizadas de forma poco habitual. Allí se encuentran tres de las actividades analizadas: Realizar pruebas sobre conocimientos previos antes de cada tema, Realizar pruebas orales, y Tener en consideración la propia autoevaluación del alumnado. En los tres casos, la puntuación modal estaría en la categoría "algunas veces" y en todos ellos hay más docentes que afirman que no las han desarrollado "nunca" o "algunas veces" que aquellos que dicen desarrollarlas "casi siempre" o "siempre". Concretamente, el porcentaje de docentes que afirma Desarrollar pruebas sobre conocimientos previos antes de cada tema es del 63,9%, realizar exámenes orales del 75,3%, y Valorar la autoevaluación del alumnado un 55,0%.

Por último se encuentra la actividad Realizar controles escritos trimestrales, en la que el 45,6% de los maestros y maestras

dicen no realizarla nunca. Esto es interesante porque es la actividad que se había encontrado que tenía un efecto negativo sobre el aprendizaje de los estudiantes.

6.2. INVESTIGACIÓN DE LAS CONCEPCIONES. SUBESTUDIO 1: CONCEPCIONES DOCENTES SOBRE EVALUACIÓN

El subestudio 1 de esta investigación tiene como objetivo determinar la incidencia del contexto (favorable o desafiante) en las concepciones docentes sobre evaluación de los estudiantes. Un primer resultado llamativo del mismo es el relativo a fuentes en las que se inspiran los docentes para hacer las pruebas. En las escuelas en contextos favorables es muy habitual usar referentes externos, esencialmente el libro de texto como base mínima y, a partir de ahí, subir más el nivel con lo trabajado en el aula. Veámoslo con tres ejemplos de tres profesores de centros diferentes:

Para elaborar la prueba me baso en el libro de texto; bueno ellos tienen libro de texto, en nuestro caso tenemos dos libros de texto de la misma editorial [...] Les pregunto un poco de eso que lo que hemos visto en clase. [P4, pri,pr,des, p.2]

Lo tenía preparado parte del año pasado, pero es que nunca doy todo igual, es decir, ciertas cosas tengo que cambiar porque a lo mejor cierto tipo de información pues he visto que no se lo he dado tan bien o he visto que ellos no lo han podido captar tan bien y entonces no insisto tanto en esto [...]. Por eso, además de basarme en el libro de texto, ahora hay material buenísimo para los profesores. Tenemos material de ampliación, tenemos material de refuerzo, y eso lo uso para los exámenes. [P8,sec,pr,fav,p.2]

Hay preguntas que tanto se han trabajado en clase con el libro como otras preguntas que han surgido en clase pues por

ejemplo, de este control que era una hoja que no sale en el tema del libro, pero sí que trabajamos en clase.
[P5,pri,pu,fav,p.2]

Los y las docentes de centros en contextos desafiantes, sin embargo, utilizan fuentes más internas, más autorreferenciales que externas. En sus discursos aparecen constantemente la elaboración mediante el trabajo en equipo y la colaboración con otros docentes de ese curso para determinar los contenidos o destrezas mínimas hasta donde quieren que lleguen sus estudiantes; curiosamente ninguno nombro el libro de texto. Como comentan dos docentes:

Yo coordino las pruebas de evaluación con mi compi del otro segundo y elaboramos juntas todo. O sea y entre otras cosas los exámenes. [P1,pri,pu,des,p.3]

Los exámenes los suelo elaborar yo y en muchas ocasiones en equipo, en colaboración, y suelo hacer exámenes muy variados y muy diferentes uno de otro. [P3,pri,pu,des,p.1]

Aunque la práctica totalidad de los docentes afirma usar una amplia variedad de instrumentos de obtención de información para la evaluación (escalas de actitudes, hojas de observación, diarios de clase...), el examen escrito, con preguntas de diferentes tipos sigue siendo tanto la prueba que más afirman usar los docentes, especialmente en Secundaria. Aun con esa pauta general, sí se observan algunas diferencias en función del contexto favorable-desafiante donde se sitúan los centros educativos.

Efectivamente, las informaciones apuntan a que los docentes que trabajan en centros cuyo contexto es favorable utilizan con mayor frecuencia y con más peso en la nota final pruebas cuantitativas. La concepción que tienen es que este tipo de

pruebas son mas objetivas y posibilitan tener información más clara a los estudiantes para mejorar su aprendizaje. Ilustrando esta idea:

Utilizo dos exámenes por trimestre. Otra prueba de evaluación es la revisión del cuaderno de trabajo de clase, pero eso solo sirve para redondear la nota. El peso de la nota lo tienen los exámenes escritos que realizan de los diferentes temas. [P9,sec,pu,des,p.3]

Si tengo que ser sincera, es eso; los instrumentos de evaluación sobre todo son: exámenes como te estoy comentando, esas pruebas orales, pequeñas exposiciones o a veces les digo, como todavía no tienen capacidad de hablar, porque hasta tercero de la E.S.O. y son muy poquitas horas [...]y ahí les evalúo y les pongo nota de su participación y demás [...] pero lo que más cuenta es el examen. [P10,sec,pri,fav,p.2]

Los docentes de escuelas en contextos desafiantes, aún usando también el examen escrito tradicional, muestran una mayor interés por el uso de pruebas "alternativas", que no muestren sólo la capacidad de los estudiantes de memorizar y redactar. Su preocupación es que deben también considerarse las diferentes habilidades y capacidades de sus estudiantes, el aprendizaje real de los estudiantes y el esfuerzo que realizan. Sirven tres ejemplos para ilustrar:

Aparte del control escrito, también evalúo la expresión oral, yo no hago en un día el control de la expresión oral, lo que pasa que sí que continuamente las actividades de clase un día le toca exponer a unos, otro día les toca exponer a otros, yo tengo mi hoja de observación que sin decir a ellos "yo te estoy evaluando" pero sí que yo le pongo cómo lo va haciendo, con la fecha y así sucesivamente otros días yo les voy preguntando [...]Otra cosa aparte es lo que apunto yo en mi diario de algo que me haya llamado más la atención [...]Los controles escritos cuentan en mi evaluación el 50%

de la nota, el otro 50% son otro tipo de competencias que considero esenciales y que no pueden medirse con un examen.
[P4,pri,pr,des,p.3]

Para mí lo más importante es el día a día de los chicos y su proceso de aprendizaje, que ellos sean conscientes. Por ello, el plan de trabajo tiene un 60 por ciento del porcentaje de la nota y los exámenes el otro 30 por ciento.
[P9,sec,pu,des,p.3]

...No suelo hacer exámenes como tal, lo veo de una manera de ponerle nota a un niño pero no se demuestran tanto sus habilidades.. [...] Procuro hacer ejercicios de evaluación en donde no se den cuenta de que están siendo evaluados y no como una prueba real, por lo que les motiva a realizarla de una manera mucho más adecuada... [P3,pri,pr,des,p.3]

Muy relacionado con el anterior elemento es el hecho de que los profesores y profesoras en contextos desafiantes muestran una especial preocupación por adaptar las pruebas a situaciones con las que se puedan sentir identificados los estudiantes. Unas palabras que reflejan este planteamiento son las siguientes:

En una prueba, por ejemplo, redacté yo como una lectura de unos niños que eran de otro barrio y que iban a una escuela por la tarde y tal, que no eran ellos pero bueno para que ellos se sintieran así, más motivados y tal. Y claro, salió bastante bien, con eso se quedaron en seguida. [P4,pri,pr,des,p.2]

Los docentes en contextos desafiantes insisten en que la evaluación sea individualizada, adaptada a las necesidades y características de cada uno de los estudiantes. La ratio por aula, las evaluaciones externas, el tiempo del que disponen o la Administración son algunos de los motivos señalados por los docentes que más dificultan lograr una verdadera evaluación individual y ajustada a cada estudiante. En palabras de los propios docentes:

Adaptar más las pruebas a cada alumno, porque realmente es muy difícil hacer un examen por ejemplo escrito, un control escrito para todos igual porque es muy difícil. Tú vas a la media, al final a la media, a los objetivos medios pero hay alumnos que para ellos es muy fácil y para otros muy difícil, y al final yo hago el control y evidentemente para los alumnos de integración y los alumnos de compensatoria, no es el mismo control. [P2,pri,pu,des,p.7]

Para mí una evaluación ideal sería una evaluación individualizada, tener en cuenta todo lo que rodea al alumno, su implicación, su esfuerzo... como te decía, no sólo la prueba objetiva como es un examen. Es que cada niño es un mundo, cada niño, cada familia que hay detrás, entonces eso, valorar todo eso. [P1,pri,pu,des,p.4]

Tener en cuenta las necesidades del alumno, sus capacidades, todo, hasta su entorno, tener en cuenta su situación familiar, todo. Yo creo que para evaluar a un alumno y sobretodo en una etapa de primaria y cuando son más pequeñitos, hay que tomar en cuenta no sólo los conocimientos que hayan adquirido que sí es súper importante pero también todo lo que hay a su alrededor. [P3,pri,pr,des,p.4]

También se observa que, frente a los que imparten clase en centros favorecidos, los docentes de centros educativos ubicados en contextos desafiantes determinan como esencial enfocar la evaluación hacia un proceso más cualitativo. Los profesores consideran que las pruebas escritas no ofrecen una información real del aprendizaje de los estudiantes, y que es necesario conocer más instrumentos cualitativos de evaluación y llevarlos a cabo, para promover un proceso más continuo de seguimiento de los alumnos, pudiéndoles dar un *feedback* o retroalimentación mucho más continua y centrada en las necesidades de cada estudiante. Así lo expresan los siguientes docentes:

Y de alguna otra manera que a mí me gustaría que fuera más oral, que fuera más individual, necesitaría una hora para cada alumno [...] habría que hacer cada vez los exámenes más abiertos y no tan cerrados, no con respuestas tan cerradas... pienso que las podría hacer un poco más abiertas... eso, que podría adaptarlas más al nivel de cada uno específico y bueno, el objetivo sería que no existieran las pruebas escritas. [P4,pri,pr,des,p.5]

Y luego me parece más justo cuando hay diferentes tipos de preguntas o diferentes formas de demostrar lo que sabes... no me gustan los exámenes... solo los exámenes, no me gustan... pero me gusta que... que los alumnos... aprendan a demostrar lo que saben poniéndolo en la práctica y dándoselo a los demás. [P9,sec,pu,des,p.4]

Por último, los docentes de centros desafiantes consideran la necesidad de que la evaluación sea un proceso constante y que pueda repetirse, dando diferentes oportunidades a los estudiantes para demostrar su conocimiento y aprendizaje. Una de las profesoras lo expresa de la siguiente forma:

También es justo que se den oportunidades, que se dé más de una oportunidad porque no puede haber una sola oportunidad, eso no es justo, porque puedes tener mal día, porque puedes [P9,sec,pu,des,p.3]

Ligada a esta idea de individualización (o no) de la evaluación, los docentes de zonas favorecidas asumen que la devolución de los resultados se realiza de forma grupal, considerando que el proceso es más transparente. Un profesor en un centro situado en un entorno favorable, indica que:

Yo les corrijo primero el examen, lo que pasa es que siempre quiero que lo corrijan y que hagan el intento de buscar qué es lo que hicieron mal y tal, entonces sí que dedico una clase para eso. [...]. Lo que suelo hacer es entregarles el control, primero lo corregimos de forma grupal para resolver las dudas. Realmente lo que les hago es, les entrego el control, en

él viene la nota y un comentario sobre qué me ha parecido su control y cómo puede mejorar, qué fallos ha tenido... y luego hago una corrección grupal [...]. En esa corrección grupal preguntan “¿y esto por qué está mal o...?”, sí, suelen preguntar cosillas, sí. [P5,pri,pu,fav,p.3]

Los docentes de contextos desafiantes, por su parte, se preocupan en que dicha devolución sea lo más individualizada posible. La idea que subyace es la necesidad de que los estudiantes puedan mejorar su aprendizaje con directrices claras:

... lo que les hago es, les entrego el control, en él viene la nota y un comentario sobre qué me ha parecido su control y cómo puede mejorar, qué fallos ha tenido... y luego hago una corrección grupal. [P1,pri,pu,des,p.2]

En esta misma línea, dos docentes nos comentan que:

Si puedo con todos lo comento uno a uno individual. Y si no, cada uno tiene señalado... no quiere decir mal, menos las tipo test, las otras yo no les tacho que está mal ni nada, por ejemplo si hay una falta de ortografía, si es una expresión que no está bien hecha o si es algo en que se han equivocado de que de eso no va la lectura, pues entonces les pongo como una flechita y ellos ya saben que ahí tienen que observar que algo pasa. Y entonces pues se comenta en voz alta, “pues yo ya me he dado cuenta de lo que es. [P4,pri,pr,des,p.4]

Evaluar es el proceso que ha tenido el alumno, detectar sus necesidades, ver si la metodología ha sido correcta, si ha avanzado o si ha retrocedido, pero no se trata de coger siete notas, hacer una media... eso no puede ser evaluar. [P2,pri,pu,des,p.5]

De los elementos resaltados en este estudio es posible configurar a la concepción que tienen los docentes de qué es una evaluación "justa" para los estudiantes. Mientras que para los docentes de centros favorables ese concepto está ligado a la imparcialidad, a la idea de una evaluación objetiva, igual para

todos. Para los de contextos desafiantes está más relacionado con la idea de valorar el esfuerzo de cada estudiante; de ahí también la importancia de la adaptación individual de las evaluaciones.

Efectivamente, para los de contextos favorables la mejor prueba es aquella que no tiene margen de error. Es decir, que es absolutamente cerrada y que no admite discusión o debate. En esta línea, un profesor indica que:

Yo lo que busco es la objetividad [...]. Corrijo lo más objetivamente posible. Mira, la de verdadero y falso es muy fácil, 25 por 0,4. [...] Me gusta de tal manera que esté estructurado y que ellos sepan... en general no suelo tener protestas como "¿y por qué me pones aquí un 1,25 sobre 2?" y le digo "a ver, ¿cuántas causas había?" y efectivamente... la ortografía y la expresión, eso va aparte. Corrijo las 25, o las preguntas que haya, si alguna pregunta vale menos porque es más sueltita, aparece también en el examen. [P8,sec,pr,fav.p.3]

La objetividad y la transparencia son dos criterios que debe tener una evaluación justa para los docentes en contexto favorecidos:

Yo ... muchas veces... yo utilizaría más actividades complementarias y utilizaría más otro tipo de trabajo... pero prima la objetividad. Te lo pide el sistema e incluso el propio alumnado. [P8, sec,pr,fav.p.4]

Creo que un componente importante de la justicia es la honestidad, entonces yo creo que tiene que haber una claridad en la evaluación y saber a qué atenerte; tener unos criterios establecidos y en función de esos criterios la persona que es evaluada sabe que va a ser evaluada y sabe para qué es la evaluación. [P5,pri,pu,fav,p.3]

[...] Hay un problema que son los objetivos... tú te tienes que ceñir a unos objetivos que te marca la ley, y se supone que si el niño no llega a esos objetivos pues no puede pasar del

cinco, se supone... no puede pasar de cinco en esa evaluación
[P9,sec,pu,fav,p.5]

Los profesores y las profesoras que enseñan en escuelas situadas en entornos desafiantes, por su parte, dan más importancia a la evolución del estudiante, al avance real de los estudiantes entre prueba y prueba de evaluación:

No corrijo de forma cerrada. Entonces lo que yo intento es valorar, es que ha habido un aprendizaje, que ha habido un avance y... el que haya una mitad de pruebas cortas... es para intentar que aprueben porque la segunda parte es mucho más difícil para los alumnos... [P9,sec,pu,fav,p.4]

Lo importante no es la nota que saquen en el examen escrito ni si avanzan puestos en el cálculo mental, sino saber de dónde ha partido y hasta dónde ha llegado. No puede ser igual de un niño que parte con una nota de un 8 y llega a un 8,5 que otro que parte con un 3 y llega a un 7. En el primer caso parece que tu aportación ha sido muy poca, mientras que en el segundo caso ha de reconocerse el logro conseguido. La nota que saques de un examen no me dice mucho, no me preocupa, sólo me detecta las necesidades. Lo que yo intento valorar a lo largo del año es el esfuerzo y el avance que ellos han tenido desde su punto de partida hasta su punto final.
[P2,pri,pu,des,p.3]

Un docente de un centro ubicado en un contexto favorable, afirma que:

Las pruebas se van adaptando a las necesidades que tú tengas del grupo. Ten en cuenta que es una evaluación, no es una evaluación firme. Me refiero que según las necesidades del grupo, la campana de Gauss que tengas, se les pide más o se les pide menos... [P4,pri,pr,des,p.4]

Por otro lado, otro profesor de escuela situada en un entorno desafiante, considera que:

Cuando ellos se dan cuenta que hay niños que les exigen más que a otros y lo preguntan “¿profè por qué a mí me pides esto, estoy y esto, y a fulanito le estás pidiendo sólo hasta

aquí?” pues le explico... Y ellos entienden perfectamente y sin decir nada, ni poner adjetivos ni calificar a nadie, ni poner etiquetas a ningún niño, hay niños que deben rendir más y hay que exigirles más.. [P2,pri,pu,des,p.2]

Complementando esta visión del contexto desafiante, una docente nos expone:

Intento adaptar las pruebas a cada niño y yo sé que es un follón con 27 niños, pero no es lo mismo hacer una prueba para diez niños que sabes que tienen más o menos el mismo nivel que para otros niños, ¿sabes? Eso no quiere decir que no les esté poniendo los mismos contenidos ni les esté poniendo diferentes objetivos, pero sí a lo mejor se los pongo de otra manera para que ellos estén más motivados. [P3,pri,pr,des,p.3]

En la misma línea, otra profesora nos expone que en una evaluación justa sería necesario:

Valorar mucho más a una persona que le costó desde el primero momento adquirir ese conocimiento y él se lo ha luchado, se lo ha peleado y ha conseguido adquirir ese conocimiento... o sea para mí no puede tener la misma nota el niño que me ha sacado las mismas cinco operaciones que desde el primer momento las coge fenomenal, que al niño que le ha costado un mes llegar a ellas. Porque a lo mejor al que le ha salido desde el primer momento fenomenal, podía haber llegado un escalón más arriba... [P6,pri,pr,des,p.4]

Asimismo, un aspecto que destacan los docentes de este tipo de contexto con más dificultades, y que les queda mucho camino por recorrer, es ser capaces de evaluar verdaderamente el esfuerzo de los estudiantes. Muchos docentes indican la dificultad de saber cuánto se esfuerza un estudiante, pero inciden en la importancia de premiar a aquellos que más se perseveran. Así, los docentes exponen que:

Siempre hay que seguir aprendiendo. Estoy aprendiendo a evaluar los esfuerzos, más que la nota en sí. Darle

importancia al trabajo diario, al esfuerzo diario, y ver de dónde parten y hasta dónde llegan. [P2,pri,py,des,p.3]

Mira una evaluación ideal sería aquella evaluación que reflejase el esfuerzo del niño durante esos 15, 20 días o un año... si tú quieres evaluar contenidos concretos o una asignatura, que seas capaz de valorar el esfuerzo que el chico ha realizado. [P6,pri,pr,des,p.4]

La que mide realmente el esfuerzo y si ese esfuerzo ha llevado a conseguir los objetivos que he planteado, tomando en cuenta de dónde parte el alumno y a dónde ha llegado, sobre todo teniendo en cuenta de dónde parte. Porque no es la cuestión llegar al mismo punto sino ver, o sea, la que evalúa la cantidad de camino que has hecho tú relacionado con tu esfuerzo. [P4,pri,pr,des,p.5]

No cabe duda que esta radical diferencia de lo que se considera una evaluación justa entre docentes de diferentes contextos socio-económicos tiene una clara repercusión en sus prácticas evaluativas y, con ello, en el desarrollo de los estudiantes.

6.3. INVESTIGACIÓN DE LAS CONCEPCIONES. SUBESTUDIO 2: CONCEPCIONES DOCENTES SOBRE UNA EVALUACIÓN JUSTA

El sub estudio 2, por su parte, busca conocer las concepciones docentes de evaluación justa. Los resultados arrojan que existen dos focos desde los cuales analizar las respuestas de los docentes: si su concepción de justicia está más cercana a la igualdad o si está más próxima a equidad. A continuación, se muestran las diferencias existentes en las concepciones de los docentes en lo que se refiere al énfasis en la igualdad o en la equidad.

6.3.1. Evaluación justa igualitaria

Los docentes que definen una evaluación justa como aquella que es igualitaria para los estudiantes destacan cuatro elementos fundamentales:

- La información transparente de la pruebas de evaluación a los estudiantes,
- La objetividad de las pruebas,
- Su carácter eminentemente cuantitativo, y
- Que les ayude a ser capaces de afrontar las evaluaciones externas a las que tendrán que hacer frente a largo de su vida.

En primer lugar, los profesores y las profesoras consideran que para que una evaluación sea justa es necesario informar a los estudiantes de cómo van a ser evaluados. Esa información se traduce en un conocimiento explícito del momento, instrumentos y criterios que el docente va a utilizar para evaluar a los estudiantes, lo que podría traducirse en una transparencia absoluta de la práctica evaluativa de los profesores. La mayoría de docentes que se posicionan en una perspectiva más igualitaria de la justicia consideran que éste es un elemento fundamental para que la evaluación de los estudiantes sea verdaderamente justa, argumentando que si éstos no saben cómo van a ser evaluados difícilmente podrán demostrar aquello que han aprendido. Mostrando estas ideas, los docentes destacan que:

Es justo que se sepa de qué te van a evaluar, qué cae en el examen, cómo va a ser la prueba, qué es lo que se va... lo que se espera que hagas para sacar buena nota...
[P8,sec,pri,fav,p.4]

Yo creo que un componente importante de la justicia es la honestidad, entonces yo creo que tienen que haber una claridad en la evaluación y saber a qué atenerse; tener unos criterios establecidos y en función de esos criterios la persona que es evaluada sabe que va a ser evaluada y sabe para qué es la evaluación. [P7,sec,pr,fav,p.3]

Este conocimiento por parte de los estudiantes ayudará a que la evaluación sea más justa, dado que los estudiantes conocerán previamente de qué, cómo y cuándo van a ser evaluados. Siguiendo esta idea, un docente indica que:

Siempre tiene que haber unos criterios de calificación expresos y los conocimientos que tienen que aprender. Tienen que estar enterados de donde están y de lo que se les está pidiendo. [P21,sec,pr,des,p.2]

En esta misma línea, otro elemento destacado por algunos de los docentes es que la evaluación tiene que ser normativa en lugar de criterial. La evaluación normativa concibe la posición de cada alumno respecto al resto de grupo-clase, pudiendo comparar y ordenar los alumnos entre ellos:

Los profesores intentamos hacer una evaluación como comparativa de todo el grupo, raras veces seguimos a un solo alumno en su tarea, entonces es bueno tener la media del grupo y ver donde se van ubicando. [D27,pri,pr,des,p.1]

En segundo lugar, los profesores y profesoras recalcan que para lograr una evaluación más justa es necesario usar pruebas cuantitativas. La concepción que tienen estos docentes es que este tipo de pruebas son más objetivas y posibilitan tener información más clara a los estudiantes para mejorar su aprendizaje. Así, éstos consideran que las pruebas o técnicas cualitativas como puede ser la observación, el diario de clase, los debates etc. no ofrecen una información real del aprendizaje de los estudiantes. Una docente afirma que:

Si tengo que ser sincera, es eso; los instrumentos de evaluación sobre todo son: exámenes como te estoy comentando, esas pruebas orales, pequeñas exposiciones o a veces les digo, como todavía no tienen capacidad de hablar, porque hasta tercero de la E.S.O. y son muy poquitas horas [...]y ahí les evalúo y les pongo nota de su participación y demás [...] pero lo que más cuenta es el examen.
[D10,sec,pr,fav,p.2]

Otro docente reflexiona que se pueden evaluar otras cosas, pero que el mayor peso de la nota de un estudiante tiene que recaer en el examen realizado:

El interés por la asignatura, expresión oral, comportamiento... puede subir hasta 1 punto la nota de las pruebas escritas, que es lo que más vale en la nota final.
[D20,sec,pr,fav,p.1]

Así, los docentes indican que el uso de pruebas objetivas, como el examen, refleja fielmente el aprendizaje de los estudiantes. En esta línea, uno de ellos indica que:

Hacerles cuatro exámenes seguidos. Lo siento pero es lo que toca. ¿Por qué? Pues porque si tú les pones tres ejercicios en un examen de un temario que tiene 20 o 25 aspectos no tengo argumentos como para poder evaluar a los alumnos. Objetivamente, tienes más probabilidades de que demuestren lo que saben. Así cuando pongo la nota, estoy seguro al 90% de que refleja lo que sabe. [D13,sec,pu,fav,p.4]

En tercer lugar, para los docentes una evaluación justa está inevitablemente ligada a la objetividad. Así, la forma más justa de evaluar a los estudiantes es aquella que no ofrece margen de error, siendo absolutamente cerrada y que no admita discusión o debate. Un profesor nos señala lo siguiente:

Yo lo que busco es la objetividad [...]. Corrijo lo más objetivamente posible. Mira, la de verdadero y falso es muy fácil, 25 por 0,4. [...] Me gusta de tal manera que esté estructurado y que ellos sepan... en general no suelo tener

protestas como “¿y por qué me pones aquí un 1,25 sobre 2?” y le digo “a ver, ¿cuántas causas había?” y efectivamente... la ortografía y la expresión, eso va aparte. Corrijo las 25, o las preguntas que haya, si alguna pregunta vale menos porque es más sueltita, aparece también en el examen.
[D8,sec,pr,fav,p.2]

Así, la objetividad y la transparencia son dos criterios que debe tener una evaluación justa:

Yo utilizaría más actividades complementarias y utilizaría más otro tipo de trabajo... pero prima la objetividad. Te lo pide el sistema, e incluso el propio alumnado.
[D8,sec,pr,fav,p.3]

Entonces, claro, tengo muy en cuenta lo que les he explicado, en lo que les he insistido y en lo que les he dicho que era más importante, trato de controlar eso porque a mí me parece que lo más importante está en la programación y bueno, en los contenidos fundamentales... y esto, que quede reflejado que el chaval lo ha asimilado, lo relaciona, lo entiende, es capaz de expresarlo bien y con la mayor objetividad posible, en cuanto a la puntuación por mi parte. [D8,sec,pr,fav,p.4]

[...] Hay un problema que son los objetivos... tú te tienes que ceñir a unos objetivos que te marca la ley, y se supone que si el niño no llega a esos objetivos pues no puede pasar del cinco, se supone... no puede pasar de cinco en esa evaluación.
[D15,sec,pu,fav,p.1]

Es muy importante ser clara con los estudiantes en cuanto a como va a ser el instrumento que les voy a aplicar, tratar de entender si lo están entendiendo o no. [D15pri,pri,fav,p.2]

Otro elemento que los docentes destacan como fundamental para llevar a cabo una evaluación objetiva del aprendizaje de los estudiantes, es el uso de rúbricas para que sepan cómo van a ser evaluados:

Tiene que ser objetiva para empezar, no puede depender del individuo, ser igual para todos. Para eso, es fundamental el

uso de rubricas, eso les orienta, saben lo que hay que hacer y además está por escrito. [D26,sec,pr,fav,p.1]

Por último, los profesores y profesoras al preguntarles qué es una evaluación justa, destacan la idea de que ésta ayude a los estudiantes a prepararles para superar las pruebas objetivas (CDI de Madrid, Pruebas de Acceso a la Universidad etc) a las que tendrán que hacer frente a lo largo de su vida académica. En este caso, la justicia es concebida en términos de preparar a los alumnos y alumnas para realizar las pruebas externas de la mejor manera posible ya que eso influirá en su vida. Así, uno de los profesores entrevistados nos indica que:

Además están ahí pendientes una serie de evaluaciones que van a ser determinantes para los chavales, para elegir sus caminos de formación y tal, y eso va a influir en cómo evaluemos también en el centro porque al final se nos va a pedir un poco desde los padres y desde la sociedad, que esos niños aprueben esas pruebas para poder seguir formándose y vivir. [D7,sec,pr,fav,p.2]

Algunos participantes en nuestro estudio tienen una visión “empresarial” de la evaluación, considerando que ésta debe preparar a los alumnos y alumnas para el día de mañana y cómo serán “evaluados” en su futuro próximo en el mundo laboral:

Mi objetivo en la evaluación es preparar a los chicos para la vida real, así que hay que intentar ser lo menos artificial posible, entonces uno de los grandes errores es evaluar de forma distinta a los estudiantes, ya que luego la vida a ellos les va a evaluar por igual. En una empresa les valoraran por igual, y hay que evitar que sientan frustración si por ejemplo no pueden graduar. Tienen que asumir sus capacidades y saber lo que pueden hacer en la vida. Este alumno a que sociedad se va a enfrentar y que conocimientos hay que potenciar y evaluar. [D31,sec,pr,des,p.2]

6.3.2. Evaluación justa equitativa

Por otro lado, existe un grupo de docentes que consideran que una evaluación justa es aquella que se basa en la equidad. Estos destacan que es necesario que:

- Dicha evaluación se adapte a las necesidades de los estudiantes y tenga en cuenta su contexto
- Mida realmente el avance o progreso en el aprendizaje
- Sea cualitativa y procesual
- Suponga un refuerzo contante y positivo para los estudiantes
- Promueva la motivación de los alumnos y alumnas

En primer lugar, los profesores definen una evaluación justa como aquella que se centra en cada uno de los estudiantes. Según los docentes, es necesario que la práctica evaluativa se adapte a las características y necesidades de los alumnos, alejándose por lo tanto de evaluaciones homogéneas. Las evaluaciones iguales para todos los estudiantes no ofrecen una visión real del aprendizaje de cada uno de ellos, por lo que no sería justo que se los evaluase a todos por igual. Los profesores consideran que cada uno de los estudiantes tiene unas capacidades y una forma de aprender y transmitir lo aprendido, por lo que, su tarea, es facilitarles la forma de enseñar lo que han aprendido. Por ello, es necesario que cada uno de ellos sea evaluado de una forma adecuada a sus características y necesidades.

Ilustrando esta idea:

La evaluación tiene que ser mucho más individualizada. Tú en concreto, David, María José o Pepito, tu evaluación es esta, por tus características concretas, por tu contexto, porque

tú has mostrado más interés, porque tú hayas sacado un 5 en un examen no te preocupes, porque tú no por ser María tienes un 5 en inglés, ¿no? Hay muchas otras cosas que hay que tener en cuenta y, en fin, hacer ver a los estudiantes que les estamos conformando como personas.

[D21,pri,pu,des,p.1]

Los docentes indican que además de ser adaptada a cada estudiante es necesario que la evaluación tenga en cuenta el contexto del estudiante así como el entorno en el que crece y se desarrolla, ya que este tiene grandes implicaciones en su desarrollo:

Pues una evaluación en la que se tengan en cuenta todo el entorno del alumno, todas sus características, su evolución, el proceso que ha tenido... [D4,pri,pr,des,p.2]

En las pruebas siempre pongo preguntas abiertas para poder adaptar la evaluación a las capacidades de cada estudiante y al esfuerzo que el alumno ha hecho. [D27,pri,pr,des,p.1]

Una evaluación pienso que debería ser lo más apegada posible a la realidad, en el sentido de captar todos los matices que pueda tener un estudiante, abarcando la mayor cantidad de habilidades posibles. [D19,pri,pr,des,p.3]

En segundo lugar, los docentes defienden que una evaluación justa es necesario que mida el esfuerzo de los alumnos y alumnas. Este esfuerzo requiere tener presente el punto de partida de los estudiantes y realizar continuas evaluaciones para comprobar su avance en el aprendizaje, conociendo la línea de aprendizaje de cada uno de ellos:

Una evaluación justa sería aquella que reflejase el esfuerzo del niño durante esos 15, 20 días o un año... Si tú quieres evaluar contenidos concretos o una asignatura, que seas capaz de valorar el esfuerzo que el chico ha realizado. Si no ha habido ningún tipo de esfuerzo porque el chico... porque los contenidos estaban poco adecuados y no ha aprendido

nada, esa evaluación está mal hecha, porque no le estás aportando nada al chico. [D2pri,pu,des,p.4]

Muy vinculada a este esfuerzo, surge entre los entrevistados la consideración de que una evaluación es justa cuando ayuda a que los estudiantes aprendan para su vida, es decir, que dicho aprendizaje sea útil y transferible a su día a día. Mostrando esta idea, un profesor indica que una evaluación es justa:

... tiene en cuenta el esfuerzo, la motivación y que hayan sacado un poco de partido. No tanto que hayan aprendido los contenidos que tengan que aprender; eso no sirve. Yo veo si un grupo o un alumno tiene resultados positivos en cuanto a si ha crecido en responsabilidad, ha crecido en un cierto gusto por el aprendizaje, o si les he despertado su curiosidad por aprender. [D22,pri,pu,fav,p.2]

Así, la evaluación tiene que contar con diferentes pruebas que realmente sirvan para la vida del estudiante, valorando aquellos conocimientos que no sean meramente teóricos y que sean aplicables a la vida real. Una docente reflexiona que:

Una prueba sola de evaluación no me sirve, ni siquiera un examen escrito. Lo importante es evaluar si saben razonar, relacionar conocimientos con otros, ver si lo que aprenden les sirve para la vida. En mis clases les enseño a pensar, a relacionar los conocimientos con otros previos que les sirvan para su vida. [D21,sec,pr,des,p.1]

Para otro de los participantes, una evaluación justa valora positivamente (premia) a aquellos estudiantes que han avanzado más en su aprendizaje, mencionando los retos como forma principal de aprendizaje y una evaluación que favorece la justicia en el aula:

Yo evalúo mucho más a una persona que le costó desde el primero momento adquirir ese conocimiento y él se lo ha luchado, se lo ha peleado y ha conseguido adquirir ese conocimiento... o sea para mí no puede tener la misma nota

el niño que me ha sacado las mismas cinco operaciones que desde el primer momento las coge fenomenal, que al niño que le ha costado un mes llegar a ellas. [P18,sec,pr,des,p.4]

Porque a lo mejor al que le ha salido desde el primer momento fenomenal, podía haber llegado un escalón más arriba... de hecho en clase lo utilizamos mucho el tema de los retos, que es una especie como de cuaderno que va semanalmente para casa en el que, los niños, te lo demandan ellos. [D3,pri,pr,des,p.2]

En la misma línea con el tema de reforzar positivamente al alumnado, otro profesor considera que la evaluación tiene que repensar el hecho de usar la nota como “premio” para los estudiantes. Así, indica que en muchos casos los docentes usan la nota como una recompensa para los alumnos, cuando el premio debería ser otro tipo de reconocimiento en el aula. El entrevistado ilustra esta idea con las siguientes palabras:

Hay que optar a que el premio no sea la nota, porque entonces empiezan las competiciones. Lo mejor es que optes por otro tipo de premios, por ejemplo, al final un alumno que lo ha hecho todo bien lo que se merece es un reconocimiento, es aumentar su responsabilidad en el aula, o una serie de cosas que con una nota al final... lo único que haces es comparar al resto, y yo busco su bienestar. [D28,pri,pu,des,p.2]

En tercer lugar, los profesores consideran que para que una evaluación sea verdaderamente justa debe alejarse de la objetividad, y contar con instrumentos y procesos más cualitativos. Así, reflexionan que es fundamental complementar los instrumentos más cuantitativos (como puede ser un examen) con información de corte más cualitativo, a través de la observación, diarios de campo o el uso de portafolios. Los docentes entrevistados reflexionan que la información cuantitativa es válida siempre y cuando se matice con instrumentos e información más cualitativa. Uno de ellos nos

explica lo que le parece esencial para alcanzar una evaluación justa:

Darle más peso a la evaluación continua, pero por otro lado considero también que es necesario hacer un examen, entonces me parece una pregunta difícil, pero sí creo que el tema de la evaluación continua es muy importante porque al final vas viendo el trabajo diario de esas personas, el esfuerzo, la progresión y al final un examen pues bueno, hay gente que funciona muy bien en los exámenes y tiene suerte.

[P9,sec,pu,des,p.2]

Algunos profesores y profesoras, indican incluso la necesidad de eliminar los exámenes como instrumento de evaluación. Así, los exámenes son vistos como elementos que afectan negativamente al aprendizaje de los estudiantes y que además no miden correctamente el conocimiento adquirido por el alumnado:

No creo en los exámenes, estresan mucho a los alumnos, no mides lo que realmente saben. Lo justo es realizar una evaluación continua por observación día a día. No me convence que se haga cada dos semanas un examen escrito, porque al final nos terminamos apoyando en el examen escrito y ahí fallamos, y los estudiantes aprenden de que lo único que cuenta es el examen y en estudiar memorísticamente para el examen, sin aprender.

[D23,pri,pu,des,p.1]

El uso no tan solo de diferentes instrumentos, sino de diferentes estrategias es mencionado por los participantes para llevar a cabo una evaluación más justa. Así, no es justo que la evaluación sea unidireccional por parte del profesor hacia los estudiantes, sino que éstos tienen que ser capaces de autoevaluar su aprendizaje, evaluar a sus otros compañeros e incluso evaluar a su docente. La autoevaluación y la coevaluación es mencionada por los profesores entrevistados

como un elemento fundamental para que la práctica evaluativa sea justa:

Es muy importante utilizar coevaluación, autoevaluación, observación, igualmente trabajos y exposiciones... no solo el examen. [D28,pri,pu,des, p.3]

Asimismo, los profesores destacan que la evaluación sea procesual, es decir, continua y formativa. Una evaluación continua es aquella que se lleva a cabo a lo largo de todo el proceso de aprendizaje del estudiante, midiendo de forma continua sus avances:

Para desarrollar una evaluación justa es importante llevar a cabo una evaluación continua pero no una evaluación continua como lo que se suele hacer de “voy poniendo positivos cada día”, sino voy viendo cómo evoluciona el alumno, si ha mejorado le valoro que ha mejorado. Una persona que estaba en un 1 si ha sacado un 3 para mí eso es una evolución o ha sacado un 5. [D16,pri,pu,des,p.2]

Haciéndole ver las diferentes evaluaciones que hay, que le evalúe un profesor, que le evalúe un compañero, o sea, que vea como que no hay diferentes clases de poder en el aula, sino que todos somos diferentes y todos podemos ayudarnos los unos a los otros tanto el profesor al alumno como el alumno al profesor y entre los compañeros también. [D5,pri,pr,fav,p.1]

Evaluar justamente es valorar el proceso que ha tenido el alumno o alumna, las necesidades ver que las has detectado, ver si la metodología ha sido correcta, si ha avanzado o si ha retrocedido, pero no se trata de coger siete notas, hacer una media... eso no puede ser evaluar. [D6,pri,pr,des,p.1]

El refuerzo positivo es otro de los elementos más importantes para los docentes para lograr una evaluación justa. Así, la información obtenida de cada estudiante a través de la evaluación es necesario que se traduzca en una retroalimentación constante y positiva, que motive al

alumnado. Reflejando esta idea, en palabras de tres entrevistados:

Decirles tú puedes mejorar, lo puedes hacer mejor, seguro que vas a llegar en la segunda evaluación; es importante explicarlo, hacer informes para entender la evolución del niño. [D25,pri,pr,fav,p.1]

Un elemento fundamental es el refuerzo positivo. Cuando un alumno mejora mucho su nota, no es tanto subírsela sino decirle lo bien que lo ha hecho, te das cuenta de la evolución que estás teniendo, esto no es fácil de conseguir... le daría la enhorabuena e intentaría que mantuviese eso. [D11,sec,pr,des,p.2]

Les doy feedback constantemente a mis alumnos, unas cinco veces cada trimestre. Les hago un gráfico individual de cada uno de los criterios de evaluación que uso, indicándoles que nota tendrían de cada criterio, y que vean donde tienen que mejorar. Mis alumnos tienen siempre muy claro que tienen que ir haciendo para ir mejorando en la asignatura. [D18,sec,pr,des,p.2]

Por último, los profesores destacan como necesario que la evaluación para ser justa tenga la capacidad de motivar a sus alumnos. Los docentes consideran que la práctica evaluativa llevada a cabo como tradicionalmente se ha hecho provoca una desmotivación en los estudiantes, ya que solo sirve para clasificarlos y darles información negativa de aquello que han hecho de forma incorrecta o que no han aprendido adecuadamente. Estos consideran que una evaluación que sea justa debería empezar por tener la capacidad de cambiar esa “negatividad” que lleva impregnada y que fuese capaz de motivar a los alumnos a querer mejorar y seguir aprendiendo:

Las imposiciones no funcionan, funciona la motivación, yo creo que a nivel de todos los seres humanos. Si yo tengo una motivación y una ilusión por algo, a mi no me cuesta tanto hacerlo, pero como no tenga ganas entonces no lo haré nunca.

Entonces nuestra labor es darles esta motivación, intentar llegar a todos, y si están motivados se ve en la evaluación, lo ven de otra forma. [D21,sec,pr,des,p.4]

Siguiendo la misma idea, otro docente complementa que:

Bueno yo creo que la principal causa de las bajas notas es la desmotivación del alumno, que yo creo que eso es un problema increíble y que no lo tenemos nada en cuenta y que realmente está ahí. Y es que si no evalúas justamente a un niño lo puedes desmotivar fácilmente, y un niño desmotivado en la primera etapa de primaria te lo cargas. Un niño motivado tiene el éxito asegurado para toda su vida. [D3,pri,pu,des,p.3]

En definitiva, podemos afirmar que existen dos tendencias claras a la hora de identificar las concepciones acerca de que es una evaluación justa. Por un lado, un grupo de docentes conciben la evaluación justa como aquella que evalúa de la forma más igualitaria y objetiva a los estudiantes; mientras que otro grupo de profesores considera que la evaluación para ser justa debe ser adaptada y diferenciada para cada alumno o alumna y tener un carácter más cualitativo, para lograr comprender el progreso en su aprendizaje.

La figura 6.1 muestra la síntesis de los principales elementos que destacan los profesores respecto a sus concepciones acerca de que es una evaluación justa.

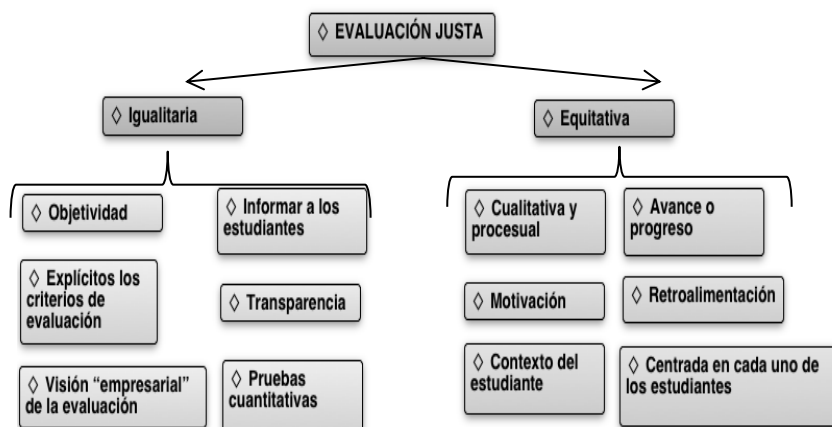


FIGURA 6.1. CONCEPCIONES DE EVALUACIÓN JUSTA DE LOS DOCENTES

Fuente: Elaboración propia.

6.4. INVESTIGACIÓN DE LAS CONCEPCIONES. SUBESTUDIO 3: CONCEPCIONES DE ESTUDIANTES SOBRE UNA EVALUACIÓN JUSTA

El análisis de las entrevistas realizadas a los estudiantes puede organizarse a través de dos grandes concepciones acerca de qué es una evaluación justa: la evaluación igualitaria y la evaluación equitativa. Algunos de los estudiantes consideran que una evaluación justa es aquella relacionada con que ésta sea igual para todos, mientras que otros estudiantes conciben la evaluación justa como aquella que se adapta a sus necesidades y características. Veámoslas.

6.4.1. Evaluación justa igualitaria

Los estudiantes que conciben la evaluación justa como igualitaria defienden que tiene que ser "igual para todos". Desde esta perspectiva, plantean la igualdad en términos de pruebas, condiciones, apoyos... Vamos a analizarlo con detenimiento.

Uno de los elementos más destacados es la necesidad de conocer las pruebas de evaluación antes de realizarla. Los estudiantes piensan que los docentes tienen que explicar cuántas pruebas de evaluación se van a llevar cabo durante el mes, trimestre, año etc., profundizando en cómo va a ser la prueba, qué contenidos va a evaluar, cuando va a ser, cuánto tiempo van a tener para realizarla y qué porcentaje de la nota final supone. Así una de las estudiantes indica que:

A lo mejor no en todos los exámenes, pero sí que nos ayudaran diciéndonos las preguntas que vayan a poner, o que en algunos exámenes pudiésemos hablar con los compañeros para poner ideas en común. [E3,sec,pr,fav,p.2]

De forma complementaria, otra estudiante nos explica que lo importante no es sólo saber cuándo y cómo va a ser el examen sino también qué espera el docente que el alumno lleve a cabo con esa prueba de evaluación. Una alumna indica que:

Para que una evaluación sea justa es importante que nos hagan exámenes y que nos expliquen cómo tenemos que responderlo [el examen], antes del examen, para saber cómo va a ser y que esperan de nosotros. [E1,pri,pu,des,p.2]

Este elemento es muy interesante, ya que los estudiantes tienen cierta incertidumbre de qué esperan los docentes de ellos en las distintas evaluaciones que llevan a cabo, así como cuál es la forma más adecuada de realizar dicha prueba de evaluación. Los alumnos y alumnas, por lo tanto, expresan la necesidad de que antes de realizar la prueba de evaluación se explicita el resultado esperado de la misma.

En esta línea, los estudiantes indican que les gustaría también que los docentes les explicasen de una forma más detallada la puntuación de las distintas preguntas o partes de cada prueba de evaluación. En el caso de las preguntas de desarrollo,

consideran indispensable que se explicita que se espera que respondan y cómo van a ser puntuados, ya que la información dada en clase o del libro de texto suele ser muy extensa. Así, demandan más claridad en cuando a la explicación y puntuación previa a la realización de la prueba de evaluación.

Otro elemento destacado por los alumnos y alumnas es la importancia de la objetividad en las pruebas de evaluación realizadas por los docentes. Muchos de los estudiantes hacen referencia a que sienten que a veces los docentes corrigen más en función del nombre de la persona que por el contenido de la prueba de evaluación. En otras palabras, los estudiantes critican la subjetividad con la que algunos docentes corrigen las pruebas, siendo necesaria la objetividad por parte del profesor. Ilustrando esta idea, un alumno indica que:

Yo creo que hay que tener en cuenta las notas, y que no miren tu nombre, es decir, que si has sacado un 5 pero te has portado mal, no te suspendan aposta yo creo que no debería mirar el nombre o el comportamiento, sino estar centrado más en tus notas, ser más objetivo. [E20,sec,pu,des,p.1]

En esta misma línea, otro estudiante reflexiona que la actitud del alumnado no se tiene que tener en cuenta, sino solamente su rendimiento en el examen:

Para que una evaluación sea justa hay que valorar a todos los alumnos por igual, dando importancia a lo que se hace, y también evitar coger manía a algunos alumnos, por ejemplo por algo que haya hecho a principio del curso, y ya ese profesor se tome todo lo que ese alumno haga bien como mal. Eso es algo injusto y aunque te haya molestado si luego el estudiante hace un buen examen hay que ponerle la buena nota, no juzgarle por algo que haya hecho un día. [E12,sec,pr,fav,p.3]

Aparte de esta idea, los alumnos y alumnas que conciben la evaluación desde una perspectiva más igualitaria consideran

que lo más justo es que la nota dependa casi exclusivamente del examen, entendido como la prueba más objetiva posible para medir su conocimiento:

Más exámenes, y menos valorar el comportamiento, porque eso es injusto. [E20,sec,pu,des,p.1]

Lo que tiene que tener más peso es el examen, porque es donde pones todos tus conocimientos. [E32,sec,pr,des,p.2]

Así, desde la perspectiva de algunos estudiantes el examen es la prueba de evaluación que refleja de una forma más fiel y justa sus aquello que han aprendido. Además, puntualizan que lo realmente justo es realizar más de un examen, es decir, hacer exámenes de forma periódica para conocer de forma constante y objetiva cómo los estudiantes van adquiriendo los conocimientos establecidos por el docente. En palabras de un estudiante:

Los docentes deberían de hacer tres exámenes por asignatura, no sólo un examen final, y tener más en cuenta el comportamiento en clase, los cuadernos, los ejercicios... pero teniendo más peso el examen, porque es más importante y muestra todos tus conocimientos. [E18,pri,pu,fav,p.3]

Además, este examen o prueba de evaluación tiene que ser la misma para todos. Los estudiantes conciben esta igualdad, tanto en lo que se refiere al tipo de prueba, como al tiempo de realización y a todas las condiciones de aplicación de la misma. Así, un estudiante considera que:

Todos somos iguales pero hay algunos que les cuesta más, y ellos se lo tienen que trabajar más o poner a todos los que les cuesta juntos en una clase, es injusto que tengan pruebas o exámenes diferentes. [E10,sec,pu,fav,p.2]

Los alumnos y alumnas añaden que es justo que todos los estudiantes realicen la misma prueba ya que si no habría estudiantes que tendrían ciertas facilidades o privilegios. Los

estudiantes piensan que si todos tienen que llegar a los mismos objetivos de aprendizaje es necesario que sean evaluados de la misma forma. Las siguientes citas muestran sus concepciones acerca de la necesidad de igualdad en el diseño y realización de las pruebas de evaluación:

Lo justo es que todos tengamos el mismo examen, y si alguno tiene menos capacidades que se esfuerce más o si no, que les separen en una clase aparte y que entre ellos hagan otro examen para todos ellos igual. [E8,sec,pr,des,p.1]

No es justo que se haga diferentes pruebas a los alumnos, la prueba tiene que ser igual para todos, porque a todos nos cuesta más algunas cosas, a mí me cuesta las matemáticas y yo me esfuerzo, y si yo lo hago, los otros se pueden esforzar en otras cosas también. [E21,sec,pu,des,p.3]

En síntesis, los estudiantes que tienen una perspectiva igualitaria de la evaluación consideran que para que sea justa, es necesario:

- Lograr que las pruebas sean objetivas, principalmente exámenes, ya que es la forma más imparcial de medir el aprendizaje y los conocimientos adquiridos por cada estudiante,
- Conocer de antemano las condiciones, características, forma de puntuación... de la prueba de evaluación,
- Conocer qué esperan los docentes de los estudiantes con esa prueba de evaluación (expectativas), y
- Promover la igualdad de condiciones en la aplicación de las pruebas de evaluación, ya que los objetivos a los que tienen que llegar son comunes en todos ellos por el currículum y eso significa que las pruebas usadas para evaluar tienen que ser también iguales.

6.4.2. Evaluación justa equitativa

Una vez analizada la perspectiva igualitaria de la evaluación, a continuación analizamos las concepciones de los estudiantes que tienen una perspectiva de la evaluación justa equitativa. En general, los estudiantes entrevistados se muestran más cercanos a este tipo de concepciones que conciben la evaluación justa como aquella que es equitativa, frente a una perspectiva más próxima a la igualdad presentada en el epígrafe anterior. Las respuestas de los estudiantes acerca de que es una evaluación justa se agrupan en siete grandes temas:

- Participación en la evaluación.
- Evaluación del progreso y esfuerzo de los estudiantes
- Uso de distintas pruebas de evaluación, complementando las pruebas cuantitativas como puede ser un examen con otras pruebas de corte más cualitativo.
- Adaptación de las pruebas de evaluación.
- Evaluación del desarrollo integral de los estudiantes.
- Evaluación interdisciplinar.
- Retroalimentación positiva al devolver la información recogida en la evaluación.

En primer lugar, uno de elementos más destacado por los estudiantes que conciben la evaluación justa desde una perspectiva equitativa e inclusiva es la necesidad de que la evaluación no se plantee de forma unidireccional docente-alumno, sino que ellos sean elementos activos en el proceso. Los alumnos y alumnas consideran que es necesario participar en la evaluación para sentir que es parte de su aprendizaje.

Por un lado, consideran que la participación e implicación en clase es un elemento clave, especialmente en la evaluación. Uno de los estudiantes indica que:

Es fundamental que nos dejen participar en clase, que nos enseñen a participar más en la evaluación.
[E9,sec,pu,des,p.1]

Con la misma idea, otra alumna puntualiza que la participación es un elemento clave para la mejora de su aprendizaje, ya que:

Si participamos, lo que tengamos que estudiar nos queda más, aprendemos más. [E1,pri,pu,des,p.1]

Por otro lado, los estudiantes también hablan de participación refiriéndose a que el docente les incluya en el diseño y aplicación de las pruebas de evaluación, ya que consideran que así sienten que tienen más responsabilidad y ello les ayuda a implicarse más en su aprendizaje:

Una vez en “Sociales” elegimos preguntas y las pusieron y eso estuvo bien, fue más justo para nosotros.
[E2,pri,pu,fav,p.2]

Ilustrando esta misma idea, otro entrevistado comparte la necesidad de tomar decisiones en cómo va a ser evaluado, creando por ejemplo, dos modelos de examen, donde el estudiante sea sujeto activo en la selección de qué tipo de prueba se adapta más a sus necesidades y capacidades:

En la evaluación nos gusta dar nuestra opinión, elegir. Por ejemplo, sería justo que nos pongan dos tipos de exámenes, por ejemplo uno tipo test y otro de explicar, porque cada uno tiene una forma de estudiar, de aprender y en la evaluación también deberíamos poder mostrar de diferentes formas lo que hemos aprendido y decidir cuál preferimos hacer.
[E9,sec,pu,des,p.3]

Los estudiantes también destacan que para que una evaluación sea justa, es necesario que sirva para ver el progreso así como el

esfuerzo que han llevado a cabo durante el proceso de aprendizaje.

Respecto al progreso, los alumnos consideran que una evaluación verdaderamente justa es aquella que tiene en cuenta si el estudiante progresa, independientemente que ese estudiante llegue o no al “5 o al aprobado”. En palabras de uno de los estudiantes:

Si los alumnos se esfuerzan, ha estado en clase atento, su actitud es buena, tiene el cuaderno bien, aunque esté suspenso, si ha avanzado mucho, yo le aprobaría.
[E9,sec,pu,des,p.2]

Otro estudiante puntualiza que el progreso es el elemento que siempre tiene que matizar la nota obtenida en el examen, ya que el examen al ser puntual es muy injusto como evaluación única. Así, las pruebas de evaluación tienen que permitir que los estudiantes muestren como han avanzado en su aprendizaje:

El examen, en una hora, no puede reflejar 9 meses de curso, por lo que los exámenes son para comprobar, pero el profesor tiene que ver el progreso de cada chaval para poner las notas. [E17,pri,pu,fav,p.2]

En cuanto progreso, los alumnos y alumnas piensan que otro elemento muy ligado al este que hay que tener en cuenta es el esfuerzo. Así, la evaluación debería tener más en consideración la implicación del estudiante y su esfuerzo que las notas obtenidas en las pruebas de evaluación llevadas a cabo. Desde su perspectiva, calificar a los estudiantes con un “aprobado o suspenso” siempre tiene que depender del esfuerzo de cada uno de ellos. Así, un alumno que no se esfuerza podría suspender de forma justificada mientras que uno que se esfuerza mucho debería aprobar aún sin llegar al mínimo exigido por el currículum:

Si por ejemplo te quedas en 4, pero te esfuerzas, es justo ponerte un 5, aprobarte, porque has trabajado mucho, y con un examen que te vaya mal porque puedes tener un mal día no es justo cargártelo todo y suspender. [E6,pri,pr,des,p.2]

La evaluación para ser más justa, tendría que tener en cuenta el esfuerzo. Si por ejemplo sacas un 4, pero te has esforzado, pues tenerlo en cuenta, y aprobarte, siempre que se vea el progreso, eso es lo más importante y muchos profesores no lo hacen. [E8,sec,pr,fav,p.4]

Además, este esfuerzo tiene que medirse y evaluarse en el día a día en el aula, viendo cómo trabaja cada estudiante, su implicación, su empeño en hacer las cosas de forma adecuada... De esta forma, si se tiene en cuenta el progreso que el estudiante ha mostrado en el aula, la evaluación sería justa ya que reflejaría su trabajo constante:

Yo creo que se tendría que tener en cuenta el trabajo del día a día, viendo el esfuerzo, porque por ejemplo, muchas veces me parece muy injusto que cuando tu has estado trabajando mucho durante el curso, y lo has estado haciendo todo lo mejor que puedes, y de repente, a lo mejor, tienes un mal examen, o un mal día suspendas. Hay algunos profesores que tiran por la borda todo lo que hemos aprendido y mejorado y eso es muy injusto. [E13,sec,pr,fav,p.2]

Los alumnos y alumnas también demandan que ésta evaluación sea continua. Una evaluación al final del trimestre no permite que los estudiantes reflejen todo aquello que han aprendido a lo largo de esos tres meses y de su trabajo diario en clase. Esta idea se ve claramente reflejada en una de las entrevistas realizadas a uno de los estudiantes:

Una evaluación justa sería una evaluación constante, donde el día a día repercuta más en la nota que un solo examen, porque no sirve de nada que durante la mayoría de días no hagas nada y no te evalúen por eso y que te lo juegues todo en un rato en un examen y eso si cuente para la evaluación. [E31,sec,pr,des,p.2]

Asimismo, los estudiantes piensan que para lograr una evaluación más justa desde una perspectiva equitativa, es necesario que el profesor realice, no una única prueba de evaluación, como podría ser un examen, sino que se usen distintas pruebas que aporten diferente información sobre el aprendizaje de los estudiantes. Los alumnos y alumnas consideran que el examen no aporta la información suficiente como para poder calificar el aprendizaje de una asignatura, de un trimestre o de un curso, ya que creen que ellos saben más que lo que se puede demostrar en un examen puntual que realizan en una hora.

Así, los alumnos y alumnas nos explican que:

No solo hay que hacer exámenes, sino otras pruebas, porque hay asignaturas en las que sólo tienes que memorizar y realmente yo creo que eso no te sirve de nada porque a las dos semanas se te ha olvidado todo. Entonces aparte que un examen, porque tienes que tener en cuenta unos datos, tienes que usar otras pruebas para conocernos mejor cómo aprendemos los estudiantes. Estas pruebas podrían ser trabajos individuales donde argumentemos o reflexionemos sobre algún tema importante. [E13,sec,pr,fav,p.1]

Evaluar con exámenes está bien, los trabajos los pondría más para subir de nota [...] los exámenes tienen que ser 70% de estudiar, memorizar, y el 30% de tu opinión, como pensamos, que se vea cómo somos, no solo que estudiamos de memoria. [E21,sec,pu,des,p.3]

Además, puntualizan que llevar a cabo un único examen es el reflejo de una situación casi puntual, de las habilidades en un momento determinado, que elude muchas otras que posee el estudiante. Además, llevar a cabo un único examen no tiene en cuenta que el estudiante pueda tener un mal día, no estar concentrado o no encontrarse bien, y eso repercutiría negativamente a su nota. Por ello, tener en cuenta solamente

un examen para la evaluación no es un fiel reflejo de lo que sabe y ha trabajado el alumno o la alumna. Para poder tener en cuenta este aspecto es importante evaluar el trabajo diario que lleva a cabo el alumnado:

Para que una evaluación sea justa también tendrían que contar otras cosas, no solo el examen, porque si te sale mal el examen ya estas suspendido...Sería interesante que cuenten la actitud, los deberes, si traes el material y esas cosas.
[E3,sec,pr,fav,p.1]

No solamente tendrían que contar los exámenes, porque en los exámenes, vale si ves si ha estudiado el chaval, pero por ejemplo no tiene por qué ser lo que haces tú durante todo el trimestre, porque por ejemplo, habrá gente que se pondrá muy nerviosa, o tendrá un mal día [...], así que también evaluaría el cuaderno, donde se ve el trabajo diario, y la actitud y la opinión, porque no es lo mismo un chaval que no hace nada, que es un mueble, que otro que participe y que aporte algo a la clase también. [E9,sec,pu,des,p.3]

Para que una evaluación sea justa hay que mirar si en clase estás atenta, si entiendes lo que explica el profesor y eso debería dar puntos, no solo el examen, que te pone nervioso o que tenías un mal día. También haría trabajos en grupo.
[E15,pri,pu,fav,p.1]

Una evaluación justa tiene que tener en cuenta todos los campos, tanto comportamiento, como examen, como trabajo en clase, como trabajo en casa, como estudio que se refleja en el examen; porque hay veces que solo te evalúan el examen o solo te evalúan el comportamiento y yo echo de menos que el examen cuente menos que el trabajo diario.
[E11,sec,pr,fav,p.1]

Los alumnos y alumnas, por lo tanto, piensan que es muy importante usar diferentes instrumentos de evaluación, ya que favorece que puedan demostrar diferentes capacidades y habilidades que en un examen no pueden demostrar. Ellos consideran que el uso de pruebas más cualitativas ayudaría a

mejorar sustancialmente su aprendizaje. Estas pruebas más cualitativas son los trabajos en grupo, los exámenes con preguntas abiertas, los debates... En esta línea, cuatro de los estudiantes entrevistados consideran que:

Se tienen que evaluar los trabajos en equipo, los trabajos en grupo, los exámenes orales, la opinión del profesor y del alumno, de la familia... y eso se tendría que tener un poco más en cuenta, porque los exámenes cuentan más y debería contar más lo que hacemos nosotros en clase.
[E28,pri,pr,des,p.1]

Una evaluación justa tendría que tener en cuenta la atención en clase, nota de exámenes y comportamiento. También se pueden hacer trabajos, deberes. Tendría que tener más peso la atención y todo lo otro porque imagínate que te sale mal el examen... y no tienes porque suspender por eso.
[E16,pri,pu,fav,p.3]

Me parece que los exámenes tendrían que ser un método con el que evaluar puntualmente, pero por ejemplo, un examen te puede condicionar mucho, es una cosa en la que te lo juegas todo a un solo día, lo que te has preparado unos días antes, y eso no puede repercutirte en un trimestre o un año entero, entonces ese método no me gusta. [E14,sec,pr,fav,p.2]

Deberían mandarnos más trabajos y evaluarlos en lugar de hacer tantos exámenes, que vean más nuestro progreso [...] con los trabajos se te queda mejor que el examen, porque te ayuda a buscar información, a pensar. [E11,sec,pr,fav,p.1]

Otro de los estudiantes entrevistados añade que si las pruebas de evaluación tuviesen un carácter más lúdico serían mucho más útiles y favorecerían más su aprendizaje:

Los profesores tendrían que hacer actividades de evaluación más divertidas, como juegos, trabajos en equipo...
[E24,pri,pr,des,p.2]

La adaptación de las pruebas de evaluación es otro de los componentes destacados para lograr una evaluación justa. Los

estudiantes entrevistados consideran que es profundamente injusto que exista una misma prueba de evaluación para todos ya que todos somos diferentes. Así, mientras que a un alumno se le da bien escribir y tiene facilidad de memorización, otro prefiere reflexionar en voz alta las ideas que ha construido con sus palabras. Si hacemos un único examen teórico escrito para todos, estamos favoreciendo al primer alumno. La mayoría de los estudiantes se cuestiona si es justo ya que consideran que las pruebas se tienen que adecuar a las necesidades, nivel y capacidad de cada uno de los estudiantes:

Creo que las pruebas tendrían que ser adaptadas, no muy diferentes, pero sí que cada uno pudiese hacerlo de la forma que mejor sabe. [E14,sec,pr,fav,p.1]

Otros alumnos y alumnas complementan esta idea reflexionando que:

Me parece justo que a las personas que no tienen las mismas capacidades se les haga otro examen, sino no es justo, por ejemplo, lo mismo que hemos dado pero un poquito más fácil o las preguntas que sean más específicas y las entienda mejor. [E20,sec,pu,des,p.2]

Una evaluación es justa si hace otro examen a los niños que les cuesta más, un examen al nivel de cada uno, porque si un niño, por ejemplo, en las matemáticas tiene algún problema y no entiende lo que se está estudiando, es justo que se le ponga algo más fácil y el resto nos enfadaríamos. [E5,pri,pr,des,p.4]

Otros alumnos reflexionan que:

Lo más justo sería que hubiese una prueba individualizada para cada estudiante y que el profesor tuviese en cuenta nuestra opinión. [E19,pri,pu,fav,p.2]

Las pruebas de evaluación tendrían que adaptarse a cada persona, por ejemplo, si tiene dificultades escribiendo que le hagan el examen oral. [E16,pri,pu,fav,p.3]

Los estudiantes no muestran problemas en pensar que cada uno sea evaluado de manera diferente, sino que, por el contrario, piensan que todos son distintos y no tienen por qué necesitar la misma prueba de evaluación para demostrar lo que han aprendido.

Otros alumnos y alumnas consideran que la adaptación de la evaluación no debe referirse solo a las pruebas sino también a los refuerzos y condiciones en los que se desarrolla dicho proceso evaluativo. Así, los estudiantes piensan que los docentes deben dar más apoyos a aquellos que más lo necesitan para poder llevar a cabo la prueba establecida (examen, trabajo, etc). Así, la justicia se traduce en dar más ayuda a los que lo requieren para que puedan llevar a cabo la evaluación de forma adecuada:

Los profesores tienen que poner un nivel medio, y si hay algunos que les cuesta más pues pueden darle más tiempo para que hagan el examen o darle ayuda fuera de horario de clase o pedirle a sus padres que le ayuden si es un trabajo por ejemplo. [E12,sec,pr,fav,p.2]

Complementariamente a todas estas ideas, los estudiantes consideran que una evaluación justa desde una perspectiva equitativa tiene que centrarse en su desarrollo integral. Así no solo debe evaluar contenidos teóricos o instrumentales (como matemáticas, lengua, inglés...) sino también evaluar la actitud, la capacidad de convivir, la empatía y el respeto hacia otros compañeros. Los estudiantes piensan que evaluar estos aspectos es imprescindible y tendría que tener más peso en la evaluación que la adquisición de conocimientos puramente académicos. Así, desde la perspectiva de los alumnos y alumnas entrevistados, es más importante participar en clase, ayudar a los compañeros, respetar sus diferencias o hacer preguntas al

profesor, que aprenderse de forma memorística algún contenido concreto:

La evaluación tendría que tener en cuenta el respeto a los profesores, a lo que haces, qué es estudiar, y hay gente que no se lo toma en serio. [E7,sec,pr,fav,p.2]

Una evaluación justa, además de valorar las notas tendría que evaluar también la capacidad de convivir que tenemos unos con otros, porque no venimos aquí solo para sacar las máximas notas, sino también para hacer las cosas mejor, facilitarles la vida de nuestros compañeros. [E12,sec,pr,fav,p.1]

Más centrado en la actitud, otro de los entrevistados considera que la forma en que los estudiantes se comportan durante todos los días en clase debería estar por encima que la calificación adquirida en un examen:

Una evaluación justa tiene que estar al nivel de los estudiantes y con varias pruebas, porque si alguien hace el examen muy bien pero se porta mal con los compañeros o maltrata a sus amigos no es justo que tenga buena nota. Hay que evaluar el comportamiento y el compañerismo y si sabemos trabajar en equipo. [E32,sec,pr,des,p.1]

Otra reflexión de los estudiantes acerca de qué es una evaluación justa es la necesidad de una mayor implicación docente en la evaluación y que éstos sean también evaluados. Desde la perspectiva de los estudiantes, en la evaluación no se trata de que los docentes pidan tareas y luego pongan una nota final a los estudiantes, ya que esto difícilmente ayuda o mejora el aprendizaje. La evaluación y las tareas requeridas por los docentes tienen que ir de la mano de una mayor retroalimentación para que la evaluación ayude a los estudiantes a saber cómo y con qué mejorar. Uno de los estudiantes nos lo refleja con las siguientes palabras:

Una evaluación justa tendría que ser dejar tiempo a los alumnos para haber estudiado bien y haberlo repasado bien con el profesor, pero sobre todo que el profesor ayude a los estudiantes y se interese, y que no diga: “pues te mando 40 ejercicios y te los estudias”, sino que tendría que ayudarnos y devolvernos esa información diciéndonos qué hemos hecho bien y en qué tenemos que mejorar y de qué forma.

[E17,pri,pu,fav,p.2]

Evaluar a los docentes para que la evaluación sea verdaderamente justa y equitativa es otra de las ideas concebidas por los estudiantes. A los alumnos y alumnas no les parece justo que ellos sean constantemente evaluados en el aula y en la escuela y ellos no puedan evaluar el trabajo que lleva a cabo el profesor. Consideran que, evaluar el trabajo docente por los estudiantes también supondría una mejora de la enseñanza y por ende de su aprendizaje a la par que se sentirían más implicados y responsables. Uno de los estudiantes nos explicó en las entrevistas que:

Sería justo que si el profesor me evalúa a mí yo también pudiese evaluar a mi profesor, porque si él me dice como hago yo mi trabajo que es estudiar, también es justo que yo opine cómo es su trabajo que es enseñar. [E7,sec,pr,fav,p.1]

Además, los estudiantes piensan que la evaluación tendría que dejar de ser tan fragmentada en asignaturas y ser más global e interdisciplinar. Ellos expresan que al salir de la escuela los aprendizajes que han adquirido tendrán que conjugarlos y unirlos para su día a día. Desde esta perspectiva, se plantean por qué dar las clases y evaluar los contenidos de forma separada, o por qué evaluar las matemáticas y las ciencias sociales de forma aislada si en la vida real dichos conocimientos tendrán que utilizarse de forma conjunta por parte los alumnos y alumnas. En palabras de uno de ellos, la evaluación para ser justa tendría que:

Ver cómo progresamos en general, no cada asignatura por su lado, porque al fin y al cabo lo que aprendemos está todo relacionado y somos el mismo alumno en una materia que en otra. A las asignaturas como mates y lenguas hay que quitarles importancia, porque parece que si las suspendes se te acaba el mundo, y no, yo creo que todas tendrían que ser iguales cuando nos evalúan. [E21,sec,pu,des,p.2]

En definitiva, como podemos ver, los estudiantes conciben la evaluación justa de forma más compleja desde una concepción equitativa. En la tabla 6.5 se resumen las diferentes concepciones de evaluación justa que tienen los estudiantes.

TABLA 6.5. RESUMEN DE LAS CONCEPCIONES DE LOS ESTUDIANTES ACERCA DE UNA EVALUACIÓN JUSTA

EVALUACIÓN JUSTA IGUALITARIA	EVALUACIÓN JUSTA EQUITATIVA
Explicitar los criterios	Adaptar los criterios según las necesidades de cada estudiante
Objetividad	Subjetividad a favor de cada estudiante
Transparencia y explicitación de los criterios	Individualización de los criterios de evaluación
Rendimiento	Valorar el progreso y esfuerzo
Examen como elemento clave de evaluación	Restar peso al examen en la evaluación
Uso de pruebas cualitativas solo como complemento del examen	Uso de instrumentos cualitativos de evaluación
Evaluación de los contenidos	Evaluación de las actitudes y valores
Evaluación igual para todos los estudiantes	Evaluación adaptada a cada uno de los estudiantes
Evaluación por materias	Evaluación holística e interdisciplinar

Fuente: Elaboración propia.



CAPÍTULO 7.

Discusión

En este capítulo se ofrece un debate dialéctico entre los resultados obtenidos en esta investigación y los diferentes hallazgos encontrados por investigaciones anteriores.

Se ha organizado la discusión de los resultados en función de las dos investigaciones que conforman la presente tesis doctoral, centradas en la incidencia de la evaluación en el aprendizaje de los estudiantes y las concepciones sobre una evaluación justa.

7.1. ESTUDIO PREVIO. INCIDENCIA DE LA PRÁCTICA EVALUATIVA EN EL RENDIMIENTO DE LOS ESTUDIANTES

Los resultados obtenidos en este estudio muestran, en primer lugar, que la forma de evaluar por parte de los y las docentes de Educación Primaria en España está directamente relacionada con el desempeño académico de los estudiantes en las cuatro áreas analizadas: Competencia Lingüística, Competencia Matemática, Competencia Científica, y Competencia Social y Ciudadana. Estos resultados, por sí solos, tienen la valía de confirmar -con datos actuales, en nuestra realidad y con una muestra estadísticamente representativa para todo el país- un hecho ya demostrado por la experiencia y la investigación (p.e. Black y Wiliam, 1998; Stiggins y DuFour, 2009; Sutton, 2009; Teddlie y Reynolds, 2000): el cómo son evaluados incide en cuánto aprenden los estudiantes. Este hallazgo, sin embargo, no por conocido es menos importante; subraya la necesidad de mejorar los sistemas de evaluación que el docente usa para que los estudiantes aprendan más y mejor.

El análisis detallado acerca de qué actividades de evaluación están asociadas al desempeño aporta resultados más interesantes, novedosos y útiles. De las ocho actividades evaluativas estudiadas por la Evaluación de Diagnóstico General (EDG), que desarrollan los docentes para Primaria, cinco han mostrado su relación con el desempeño en alguna competencia.

En primer lugar, se ha observado que realizar controles después de cada tema o lección incide positivamente en el aprendizaje de los estudiantes. Este resultado hace referencia a la importancia de tener la enseñanza bien estructurada en lecciones, y en las que en cada una estén perfectamente

delimitados los objetivos, los contenidos, los deberes y las actividades de evaluación de carácter formativo (Hunt, Wiseman y Touzel, 2009; Killen, 2005; Murillo, Martínez-Garrido y Hernández-Castilla, 2011).

Un docente que realiza evaluaciones al final de cada tema es un docente que hace evaluaciones frecuentes. Y la investigación ha encontrado relación entre la evaluación frecuente y el desarrollo cognitivo del estudiante (Muijs, Chapman, Collins y Armstrog, 2010; Murillo, 2007a, 2007b), pero también con el no desarrollo socio-afectivo (Kyriakides, Creemers y Antoniou, 2009; Murillo y Hernández-Castilla, 2011). En esa lógica puede también entenderse otro de los hallazgos de este estudio: los estudiantes de los docentes que realizan controles trimestrales escritos obtienen peores resultados en comparación con los que no los realizan. Seguramente porque los maestros y maestras que realizan controles al final del trimestre no realizan estas evaluaciones frecuentes de las que hablamos.

Otra actividad evaluativa que, según esta investigación, incide positivamente en el desempeño de los estudiantes, es que el docente corrija los deberes y cuadernos; al igual que revisar las actividades que hacen en clase. Estos resultados son coherentes con los encontrados por Carlson y Francis (2002). Estos investigadores encontraron que cuando los docentes corrigen las actividades realizadas por los estudiantes y se las devuelven hay un aprendizaje significativamente mayor de éstos.

Especialmente interesante es el hallazgo que establece una relación directa y medible entre la valoración del interés y la participación en clase de los estudiantes y su aprendizaje. Otro estudio anterior, de carácter experimental, realizado por Vey (2005), refleja la mejora del rendimiento de los estudiantes

cuando la evaluación considera la colaboración y comunicación entre ellos.

Esta investigación, sin embargo, no ha encontrado relación entre Realizar una prueba sobre conocimientos previos antes de cada tema, Realizar exámenes orales y Valorar la autoevaluación del alumnado, y el desarrollo de competencias por parte de los estudiantes. Estos resultados no son coincidentes con otras investigaciones desarrolladas en el ámbito internacional (ver, por ejemplo, Muijs et al., 2010). Habrá que seguir investigando las razones de esas discrepancias.

Como síntesis de los resultados obtenidos, en la figura 7.1 se han representado las relaciones encontradas entre actividades de evaluación desarrolladas y las competencias en las que impactan.

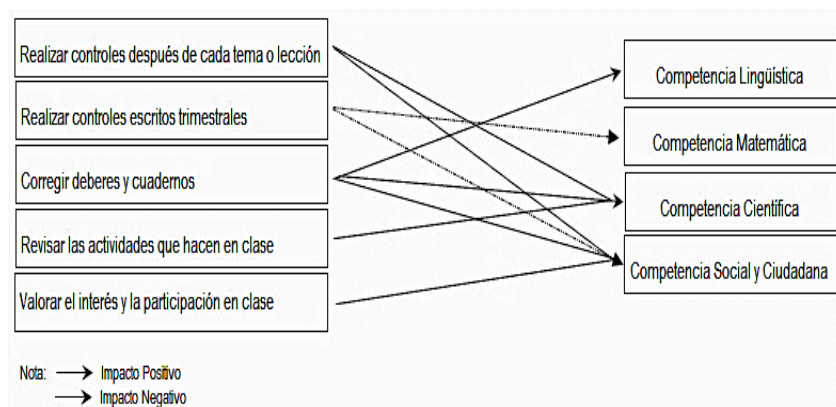


FIGURA 7.1. INTERRELACIÓN ENTRE LA ACTIVIDAD DE EVALUACIÓN QUE DESARROLLA EL DOCENTE Y LA COMPETENCIA EN LA QUE IMPACTA

Fuente. Elaboración propia.

7.2. LAS CONCEPCIONES SOBRE EVALUACIÓN

En esta investigación, en primer lugar, hemos obtenido evidencias empíricas claras de que el contexto socio-económico en el cual se desarrolla la acción docente determina las concepciones evaluativas de los profesores y profesoras. Se han encontrado diferencias en las concepciones sobre los referentes de la evaluación, sobre las pruebas utilizadas y su adaptación, sobre la devolución de información. Todo ello para llegar a configurar una concepción de “evaluación justa” radicalmente diferente en un grupo de docentes y en otro.

Si relacionamos estos hallazgos con las categorías que Brown (2003, 2004) utilizó, observamos que efectivamente el contexto hace que las concepciones de los docentes se ajusten más a una categoría que a otra. Como elementos extremos vemos que los docentes de escuelas en entornos favorables consideran que la evaluación sí es un instrumento válido para la rendición de cuentas en la escuela y que efectivamente miden los logros producidos por esta y por el profesor. Estas ideas, sin embargo, son radicalmente rechazadas por los docentes que laboran en contextos desafiantes. En el otro extremo, se observa una tendencia de los docentes de escuelas desafiantes a considerar la evaluación como un proceso irrelevante (Brown, 2003), que no aporta gran cosa y que tiene muchos elementos negativos para el desarrollo de los estudiantes. Las concepciones de evaluación como mejora o como rendición de cuentas del propio estudiante (Harris y Brown, 2008) se encuentran en ambos colectivos, pero con muchos matices. Con el planteamiento de conseguir buenos resultados, con referentes externos, o con la idea de valorar el crecimiento del estudiante. Estos hallazgos se confirman en el subestudio 2, donde las concepciones de

evaluación justa del profesorado varían en función del entorno en el que se encuentre la escuela.

El estudio de Brown, Hui, Yu y Kennedy (2011) es un referente al señalar la incidencia del contexto en las concepciones de evaluación. Allí se destaca la importancia del contexto en la evaluación de los estudiantes: la cultura, así como el estatus social de los docentes, inciden en la forma en que estos conciben la enseñanza y la evaluación. Estos autores encontraron que mientras que algunas concepciones educativas pueden ser esencialmente iguales en toda la población, las concepciones sobre la evaluación son legítimamente diferentes por las diferencias sociales y culturales existentes en las escuelas y en la sociedad.

Si se relaciona los resultados de este estudio con las dos grandes concepciones de evaluación que aborda Remesal (Coll y Remesal, 2008; Remesal, 2011), la más "pedagógica" y la más "social" (y la mixta), se observa una tendencia a tener una visión más social los docentes de contextos favorables y más pedagógica los de contextos desafiantes; eso así dicho con todas las precauciones.

También se han aportado evidencias del gran salto entre los deseos y la realidad, al menos en la evaluación de los estudiantes, tal y como encontraron Winterbottom y sus colaboradores (2008). Efectivamente, todos los docentes afirmaron que su práctica evaluativa real dista mucho de lo que ellos consideran que es una evaluación óptima del aprendizaje de los estudiantes.

Sin embargo, lo más sugerente de este estudio es la aparición del concepto de "evaluación justa" y cómo es visto de forma diferente por los docentes de contextos favorables y

desafiantes. El tema de la "justicia", entendida como "*fairness*", es uno de los temas más importantes y que ha atraído a muchos investigadores del ámbito anglosajón (Gipps y Stobart, 2009; Pettifor y Saklofske, 2012), aunque en muchas ocasiones asociado a calidad de los test y su construcción. Menos investigación hay sobre las concepciones de lo que es una evaluación justa por parte de los docentes. Estos resultados abren una interesante línea de trabajo futura sobre qué consideran los docentes y una evaluación justa y especialmente sobre los factores que hacen que se construya esa concepción y qué repercusiones tiene para su práctica. La investigación de Scott et al. (2014) es una buena orientación al respecto.

Profundizando en los hallazgos del subestudio 1, en el subestudio 2 presentado se han obtenido evidencias empíricas de las percepciones de los docentes acerca de qué es una evaluación justa. Esta línea de investigación, poco estudiada hasta el momento, aporta evidencias sobre las ideas implícitas de los docentes acerca de la evaluación y cómo se traducen estas en su forma de evaluar a los estudiantes.

Los resultados de las entrevistas realizadas a los docentes muestran que existe un elemento clave en las diferentes concepciones de los docentes sobre qué es una evaluación justa (igualitaria o equitativa): el contexto en el que se encuentra el centro educativo en el que trabajan, coincidiendo con los resultados obtenidos en el subestudio 1. Así, los docentes que trabajan en entornos favorables consideran que una evaluación justa es aquella objetiva, cuantitativa y transparente en cuanto al proceso. No obstante, para los profesores que enseñan en centros ubicados en contextos desafiantes y que tienen una diversidad de estudiantes mayor, indican que la evaluación

justa debe adaptarse a las necesidades de ellos, ser más cualitativa y medir el progreso real de los alumnos.

Otro elemento encontrado que no incide en la forma de concebir que es una evaluación justa es el nivel educativo en el que ejercen los docentes. Así, no hemos encontrado diferencias en cuanto a la concepción más igualitaria o equitativa de la evaluación por ser profesor de educación primaria o secundaria.

Si relacionamos estos hallazgos con la propuesta de Lam (1995) vemos que los resultados coinciden con los docentes que tienen una perspectiva más cercana a la equidad. Así, tanto Lam (1995) como los docentes entrevistados consideran que una evaluación es injusta si los estudiantes no tienen las mismas igualdades de oportunidades para demostrar aquello que conocen, si las evaluaciones son sesgadas y solo sirven para evaluar un tipo de estudiantes o alguna capacidad en específico y por último los docentes toman decisiones en detrimento de aquellos estudiantes que presentan más dificultades.

Los resultados de nuestra investigación también están en concordancia con Suskie (2000) la cual considera que para lograr una evaluación justa es necesario: a) establecer claramente las metas y los resultados de aprendizaje, b) unir la evaluación con aquello que se enseña y viceversa, c) utilizar distintos instrumentos y pruebas de evaluación así como evaluar en diferentes momentos, d) ayudar a los estudiantes a aprender a realizar satisfactoriamente las pruebas de evaluación; e) empoderar a los estudiantes, y f) interpretar los resultados de la evaluación de forma adecuada.

Por último, las dos concepciones de evaluación justa que tienen los estudiantes (igualitaria *versus* equitativa) encontradas en el subestudio 3, están relacionadas con las dos acepciones del

adjetivo “justo”. De un lado se asocia con la justicia legal (o general) y de otro con la justicia distributiva (o social). Efectivamente, algunos estudiantes, relacionan evaluación con el establecimiento de un juicio de valor y, como tal, debe basarse al principio de que “todos los seres humanos son iguales ante la ley, sin que existan privilegios ni prerrogativas de ninguna clase”, lo que se conoce como igualdad legal. Por otro lado, otros estudiantes tienen concepciones de evaluación justa más cercanas a la Justicia Social; es decir, basadas en el principio de redistribución, sensibilización y de apoyo diferencial (Murillo e Hidalgo, 2015b).

Estas dos ideas de justicia hunden sus raíces en Aristóteles, concretamente en la obra “Ética a Nicómaco” (Aristóteles, 350 a.C./2004). Allí sostiene que la virtud más necesaria de todas para la conservación del mundo es la virtud de la justicia, considerada como la suma de todas las virtudes. Distingue dos tipos de justicia, la justicia general, de la que deriva la justicia legal, y la justicia particular, entre las que se encuentra la distributiva, que es la base de la Justicia Social que luego retomarán Platón (380 a.C./2003), Tomás de Aquino (1485/2002), hasta llegar a Rawls (1971) y de él a nuestros días (Murillo y Hernández-Castilla, 2011).

Este binomio igualdad *versus* equidad en las concepciones de evaluación justa, no es novedoso en la literatura. Así, Tony Lam (1995) considera que garantizar la igualdad en la evaluación permite comparar los resultados y simplificar las evaluaciones realizadas, pero se pierde significación, dejando de ser imparcial; mientras que garantizar la equidad reduce el sesgo y permite una evaluación significativa pero introduce dificultades en la administración y en la comparación de los resultados de los estudiantes. Para este autor una evaluación

justa desde planteamientos igualitarios se centra en garantizar la igualdad de condiciones en la realización de la prueba (materiales, tiempo, espacio, recursos...). Una evaluación justa igualitaria requiere la misma administración, contenidos, puntuación e interpretación de los resultados, persiguiendo, por lo tanto, la objetividad.

Los hallazgos de este subestudio 3 están en coherencia con el subestudio 2 y la propuesta de evaluación justa de Suskie (2002). Los estudiantes defienden que una evaluación justa no es aquella que es igual para todos los estudiantes, ya que no todos han tenido necesariamente las mismas oportunidades de acceso a la educación; por lo que el principio de igualdad no es suficiente para garantizar la justicia. Suskie (2002) también considera esencial la promoción de la equidad en la evaluación para garantizar la justicia, adaptando la evaluación a las necesidades y características de los estudiantes. Esta equidad se traduce en llevar a cabo múltiples evaluaciones y con diferentes instrumentos para asegurar que la diversidad de los estudiantes se atiende correctamente a través de la evaluación.

Analizando los resultados del estudio de Pepper y Pathak (2008), existen algunas coincidencias con nuestra investigación. Así, los estudiantes del subestudio 3 coinciden con los autores texanos en considerar que la explicitación de los criterios de evaluación es un factor esencial para desempeñar una evaluación justa. Pepper y Pathak (2008) por su parte, también resaltan la necesidad de una retroalimentación frecuente, señalada igualmente por nuestros participantes.

Relacionando el presente estudio con la investigación llevada a cabo por Scott y colaboradores (2014), encontramos algunos elementos comunes. En coherencia con sus resultados, desde

las concepciones equitativas de una evaluación justa de los estudiantes, estos consideran que la evaluación debe adaptarse a las necesidades y características de cada uno de los estudiantes así como que sea integral. Asimismo, Scott y colaboradores (2014) remarcen la necesidad de objetividad y transparencia en el proceso evaluativo, también destacado por los estudiantes que conciben la evaluación desde una perspectiva igualitaria de nuestro estudio.

Resulta interesante el hecho de que no se hayan encontrado diferencias en las concepciones de los estudiantes en cuanto al nivel educativo, el contexto de la escuela o su titularidad. Este análisis se contrapone con los hallazgos encontrados tanto en el subestudio 1 como en el 2, donde sí existen diferencias en las concepciones del profesorado en cuanto al contexto en el que se encuentran las escuelas. Parece necesario seguir profundizando en los factores que inciden en tener una concepción igualitaria o equitativa de una evaluación justa.

En síntesis, podemos afirmar que los hallazgos relativos a la incidencia del contexto coinciden con resultados obtenidos en estudios análogos (Brown, 2003, 2004, 2008; Brown, Hui, Yu y Kennedy, 2011; Harris y Brown, 2008; entre otros). Asimismo, en los subestudios 2 y 3 encontramos que existen dos tipos de concepciones en las ideas implícitas acerca de una evaluación justa: una más próxima a la igualdad y otra más cercana a la equidad. La existencia de estos dos polos en la concepción de una evaluación justa también ha sido señalada por autores como Scott et al. (2014), Suskie (2002), Pepper y Pathak (2008) o Lam (1995) entre otros. En este sentido, nuestra investigación aporta datos del contexto español de esta reciente línea de investigación tan importante, la cual es necesario que siga

siendo estudiada si queremos mejorar la práctica evaluativa en las escuelas.



CAPÍTULO 8.

Conclusiones

En este capítulo se presentan las ideas conclusivas de esta investigación organizadas en cuatro apartados. En primer lugar, se describen las lecciones aprendidas y sus implicaciones para una evaluación de estudiantes más justa. El siguiente apartado recoge las fortalezas del estudio, así como las limitaciones encontradas a lo largo del desarrollo del mismo. Posteriormente, se sigue con una mirada a las posibles líneas de investigación que marcarán nuestro trabajo futuro. Unas ideas finales cierran este capítulo y, con ello, esta tesis doctoral (con el permiso de las referencias y unos voluminosos anexos).

8.1. LECCIONES APRENDIDAS DE LA INVESTIGACIÓN

El estudio previo nos ha legado una idea conclusiva fundamental: la importancia de la evaluación para el desarrollo integral de los estudiantes. Qué, cómo, cuándo, dónde, con quién y para quién de la evaluación son decisiones que inciden de forma directa y medible en el aprendizaje de los y las estudiantes. Una evaluación inadecuada, como hemos visto, puede incluso entorpecer el desarrollo global de los estudiantes. De ahí que sea necesario ponerlo en primera línea de atención en la formación tanto inicial como permanente de los docentes.

En este sentido, las prácticas de evaluación que más inciden en el rendimiento académico del alumnado en Educación Primaria, y que por lo tanto más se deberían potenciar en el aula, son: a) el seguimiento continuo del aprendizaje, b) revisar deberes y cuadernos y dar una retroalimentación a los estudiantes, c) valorar el interés y la participación en clase de los estudiantes, y d) considerar todas las actividades desarrolladas en clase. Estos hallazgos nos permiten afirmar que si queremos que la evaluación sea una herramienta útil y para el aprendizaje se debe orientar a estos aspectos, incluyendo, asimismo, a los estudiantes en el proceso evaluativo.

Repensar la evaluación supone avanzar hacia procesos más inclusivos, participativos e individualizados para conocer los progresos de los estudiantes, alejándonos del uso de pruebas cuantitativas únicas y dejando paso al uso de varias pruebas y técnicas de evaluación eminentemente cualitativas que realmente permitan conocer la evolución, aprendizaje y desarrollo de cada uno de los estudiantes.

Un segundo aprendizaje de esta investigación es la relevancia del estudio de las concepciones sobre evaluación para conocer y comprender en profundidad cómo evalúan los y las docentes. Investigar acerca de sus concepciones tiene un indudable valor para mejorar las prácticas docentes. Conociendo y comprendiendo cómo son y cómo se construyen esas concepciones será posible modificarlas (Remesal, 2011).

La forma en que los docentes conciben su evaluación y cómo esta se traduce en su práctica diaria son elementos clave para la mejora educativa. Ser conscientes de cómo se evalúa y qué pasos faltan para llegar a una mejor evaluación es uno de los retos del sistema educativo, así como implicar más a los estudiantes en su proceso de evaluación y de aprendizaje.

Otra de las conclusiones de esta investigación es que los docentes tienen distintas concepciones acerca de una evaluación justa. Así, sus concepciones se encuentran en un continuo de dos polos: la igualdad y la equidad. Mientras que hay docentes que consideran que una evaluación será justa solo si es igual e idéntica para todos, otros piensan que la justicia evaluativa se refiere a adaptar la evaluación a cada uno de los alumnos o alumnas y su realidad.

Uno de los aspectos detectados que más incide en las concepciones sobre evaluación (y especialmente en las ideas implícitas acerca de qué es una evaluación justa) es el contexto en el que trabajan los profesores y las profesoras. No es azaroso que las decisiones políticas actuales que fomentan de una forma desmesurada evaluaciones de estudiantes externas, incluso con consecuencias duras, tengan una visión de la evaluación más cercana a planteamientos igualitarios, y con ello, más cercana a la que tienen los docentes de escuelas situadas en contextos

favorables. De esta forma, una vez más, son los estudiantes de contextos desafiantes los que se enfrentan a evaluaciones alejadas de su práctica en el aula.

Si mantenemos planteamientos igualitarios de evaluación seguiremos favoreciendo a los que más tienen y menos necesitan. Es necesario romper con las políticas heterogéneas de evaluación y convertir la práctica evaluativa en un procedimiento equitativo, sensible con el contexto de las escuelas y con las necesidades y particularidades de cada uno de los estudiantes. Una evaluación igual para todos deja en clara desventaja a los que, por sus características, necesitan más recursos, tanto personales como materiales.

No deja de ser interesante que el contexto solo influye en las concepciones docentes. Por el contrario, en las concepciones de los estudiantes no podemos afirmar esa relación. Así, aunque tienen concepciones cercanas a la evaluación igualitaria o equitativa, estas no vienen determinadas por el contexto del centro educativo. Esto nos hace reflexionar cómo la presión política evaluativa a las escuelas, la experiencia docente y la cultura de los centros inciden en las concepciones de evaluación, claramente marcadas en los profesores entrevistados. Si queremos que los docentes evalúen de forma equitativa es fundamental que las políticas de evaluación cambien dejando de ser iguales para todos y atendiendo de forma diferencial a las escuelas que más lo necesiten, evitando comparaciones y *rankings* que nada aportan al aprendizaje y desarrollo integral de los estudiantes.

Por último, tanto el estudio previo como el trabajo de las concepciones reflejan lo mucho que hay que avanzar para que se desarrollen en el aula evaluaciones que de verdad estén en

consonancia con planteamientos actuales de un currículo que contribuyen a una educación que trabaje por la Justicia Social. A pesar de discursos, sigue habiendo prácticas evaluativas caracterizadas por tener una visión jerárquica, autoritaria y de control de los estudiantes y de su aprendizaje. De esta forma, parece urgente fomentar el desarrollo profesional de los y las docentes en estrategias alternativas de evaluación educativa tanto en la formación inicial como en la permanente. Solo con otra evaluación tendremos una educación que contribuya al cambio social.

8.2. FORTALEZAS Y LIMITACIONES DE LA INVESTIGACIÓN

La realización de esta tesis doctoral nos ha permitido detectar ciertas potencialidades del estudio, pero a su vez algunas limitaciones que es importante destacar.

En primer lugar, el uso de la fenomenografía en la investigación cualitativa es una de las fortalezas más relevantes de la presente tesis doctoral. Utilizar esta metodología para profundizar en las concepciones de cómo tiene que ser una evaluación (y especialmente una evaluación justa), ha sido muy poco utilizado en estudios españoles. En este sentido, las 73 entrevistas realizadas aportan una visión completa sobre las concepciones de docentes y alumnado de Educación Primaria y Secundaria, profundizando no en sus prácticas de aula sino en sus ideas implícitas acerca del proceso evaluativo. La mayoría de las investigaciones nacionales utilizan otras metodologías cuantitativas o cualitativas porque tienen por objetivo conocer el quehacer docente en sus aulas, mientras que el interés de

nuestro estudio recae en sus concepciones, pudiendo coincidir o no con su práctica evaluativa.

Otro aspecto novedoso de la investigación es la temática abordada. La investigación de una evaluación justa ha sido poco investigada hasta el momento desde un punto de vista empírico, por lo que nuestra investigación aporta datos útiles para el avance del conocimiento y para mejorar la educación. Si queremos mejorar la educación es fundamental empezar a desarrollar procesos evaluativos más justos, siendo relevante profundizar en las ideas implícitas de los protagonistas del proceso de enseñanza-aprendizaje.

Muy ligado a lo anterior, este trabajo no se centra únicamente en las concepciones docentes, sino también en las de los estudiantes. Las ideas implícitas del profesorado y el alumnado no siempre coinciden, aportando una visión llena de matices de cómo tendría que ser una evaluación justa y cuál es el papel que desempeñan ambos actores en ella. En este sentido, esta tesis ha contribuido a ofrecer una visión complementaria de las concepciones de docentes y estudiantes. Este aspecto, ha permitido profundizar en sus similitudes y diferencias en las ideas implícitas acerca de cómo tiene que desarrollarse la práctica evaluativa para ser justa.

Otro de los puntos fuertes de esta investigación es que da voz a los protagonistas del proceso educativo. Profundizando en las investigaciones centradas en las concepciones sobre evaluación, la mayoría de ellas utiliza un enfoque cuantitativo a partir del análisis con ecuaciones estructurales de sus ideas implícitas. Nuestro estudio aporta matices y refleja las ideas de los protagonistas sobre la evaluación. No solo nos interesa ubicar

sus concepciones en un tipo de evaluación, sino otorgar presencia a sus palabras y comprender sus reflexiones.

La investigación presentada en esta tesis doctoral, como cualquier otra, no está exenta de limitaciones. Ser consciente de los puntos fuertes pero también de aquellos aspectos que podrían mejorar la investigación es fundamental si queremos mejorar la investigación y la educación.

La primera limitación que destacamos de esta tesis es que el estudio previo se centra en Educación Primaria. En este sentido, sería interesante estudiar la incidencia de la evaluación en el aprendizaje de los estudiantes de Educación Secundaria y conocer, asimismo, qué prácticas evaluativas determinan en mayor medida su rendimiento académico.

En segundo lugar, en el estudio fenomenográfico, las concepciones ambiguas, contradictorias o cambiantes de muchos docentes hacen que no haya sido posible ubicar claramente a todos los docentes en un grupo u otro, con lo que, en segundo lugar, esta trabajo se muestra incapaz de determinar qué hace que los docentes tengan una concepción más igualitarista o más equitativa de lo que es una evaluación justa.

Por su carácter fenomenográfico, otra de las debilidades es el uso exclusivo de un enfoque cualitativo. En este sentido, sería interesante complementar el estudio cualitativo con otro cuantitativo con una muestra representativa nacional de docentes que ofreciese la posibilidad de generalizar los resultados obtenidos.

En el estudio de las concepciones, los participantes no son los mismos en los tres subestudios realizados (el primero centrado en la incidencia del contexto en las concepciones de evaluación,

el segundo en las ideas implícitas de docentes sobre una evaluación justa, y el tercero en las concepciones de estudiantes sobre una práctica evaluativa justa). Mientras que el primer subestudio cuenta con diez participantes, las concepciones sobre una evaluación justa del segundo y tercer subestudio tienen 31 docentes y 32 estudiantes que participaron en la fenomenografía. Sería bueno ampliar el primer subestudio a más docentes y también extenderlo a estudiantes.

En relación a los participantes de nuestra investigación, también hubiese sido interesante ampliar la muestra a otras Comunidades Autónomas. Por temas de accesibilidad, la investigación se realizó íntegramente con docentes y estudiantes de la Comunidad de Madrid, siendo interesante entrevistar a profesorado y alumnado de otras comunidades de España para, por un lado, confirmar los hallazgos encontrados y por otro, comparar los resultados y conocer así la existencia de diferencias entre sus ideas implícitas de la práctica evaluativa.

A pesar de las debilidades señaladas, consideramos que las fortalezas y avances que han representado este estudio aportan evidencias claras sobre la importancia del estudio de las concepciones, fundamental si queremos mejorar la educación mediante el avance hacia una evaluación más justa en las escuelas.

8.3. LÍNEAS FUTURAS DE INVESTIGACIÓN

Una vez mencionadas las principales lecciones aprendidas de esta tesis doctoral así como sus debilidades y limitaciones, nos planteamos algunas posibles líneas de investigación que permitan enriquecer los hallazgos obtenidos. Sin duda, aún

queda mucho camino por recorrer en las concepciones sobre evaluación y, desde nuestra perspectiva, la investigación podría tomar diversos rumbos que complementasen y mejorasen el estudio presentado en esta tesis.

Asumiendo las fortalezas y debilidades de la presente tesis sería interesante seguir profundizando en los aspectos más innovadores y potentes de este estudio y diseñar otras líneas de investigación que atenúen o superen las debilidades identificadas.

Una línea de investigación que puede ser abordada en futuros estudios sería seguir investigando en la incidencia del contexto socio-económico en las concepciones de los docentes. Y las conclusiones son claras: dicho contexto determina las concepciones y con ello las prácticas. De esta forma, frente al supuesto tratamiento indiferenciado por parte de las Administraciones de todos los centros, la investigación muestra que la segregación escolar configura enseñanzas muy diferentes en función del contexto socio-económico. Por ello, sería necesario realizar una investigación, tanto cuantitativa como cualitativa para determinar en qué medida incide el contexto en la práctica evaluativa de los docentes y en sus concepciones.

Tal como lo indicamos en las debilidades, en este estudio quedó pendiente complementar los datos fenomenográficos con un estudio cuantitativo ex post facto con la misma temática. Para realizarlo, sería interesante seguir la línea iniciada por Gavin T. L. Brown, adaptando y traduciendo el inventario *CoA (III)* de Brown (2003, 2004, 2006, 2009) para conocer las concepciones tanto de docentes como de estudiantes sobre evaluación. Una vez aplicado el cuestionario se analizarían las concepciones a través de Modelos de Ecuaciones Estructurales. Dicho análisis

permitiría conocer qué distintos tipos de concepciones tienen tanto docentes como estudiantes. Con los resultados de este estudio se podrían establecer relaciones con los resultados obtenidos en la presente investigación fenomenográfica y comprobar si existe coherencia entre los enfoques cuantitativos y cualitativos. En este sentido, ya se han iniciado los primeros contactos con el investigador neozelandés Gavin T. L. Brown para empezar con la traducción y validación de su inventario al español.

Asimismo, sería interesante replicar un cuestionario como el de Brown y sus colaboradores (2003, 2004, 2008, 2009) centrados en las concepciones sobre una evaluación justa de los estudiantes. Analizar las concepciones de una práctica evaluativa justa a partir de los Modelos de Ecuaciones Estructurales aportaría también una visión llena de matices que podría complementar los datos aportados en esta tesis mediante la fenomenografía.

Siguiendo otras investigaciones realizadas y con grandes implicaciones para la práctica educativa (Azis, 2014, 2015; James y Pedder, 2006; Prieto, 2008; Vandeyar y Killen, 2007; entre otros), otro estudio relevante se centraría en profundizar en la relación entre concepciones y prácticas de evaluación en un mismo estudio. Dado que las concepciones determinan la práctica evaluativa sería relevante conocer, por un lado, las ideas implícitas de un docente y, por otro, cómo evalúa en el aula a través observaciones. Los estudios que han profundizado en esta temática de investigación determinan que existe una brecha entre las concepciones y la práctica de evaluación del profesorado, siendo interesante realizarlo en el contexto español y confirmar o rebatir dichos hallazgos.

Otro elemento que merece ser estudiado es conocer la incidencia de las políticas de educación y en especial de evaluación en las concepciones sobre evaluación. Las Administraciones determinan, en gran medida, las actividades de evaluación de los centros escolares, pudiendo influir en las ideas implícitas del profesorado acerca de su práctica evaluativa, y en especial, de una práctica evaluativa justa. Por ello, sería fundamental seguir investigando en esta línea y aportar datos relevantes que pusiesen de relieve las repercusiones de las pruebas externas en la práctica diaria evaluativa de los y las docentes. En este sentido, sería interesante analizar las políticas de evaluación de la Administración, realizando un estudio *ex post facto* o cualitativo con los responsables nacionales y locales de evaluación. Una vez conocidas sus prácticas y sus concepciones de evaluación éstas se podrían comparar con las obtenidas en las escuelas y complementarlas con un estudio cualitativo para conocer la ideas implícitas del profesorado acerca de cómo les influye las evaluaciones externas a su práctica diaria.

No cabe duda que es urgente seguir investigando sobre la evaluación para la Justicia Social. En este sentido, y tal como se ha expuesto en apartados anteriores de esta tesis doctoral, paralelamente al estudio sobre las concepciones de evaluación estamos profundizando en estudios teóricos acerca de cómo tiene que ser una evaluación para promover el cambio social. Estos avances teóricos sobre la evaluación para la Justicia Social son necesarios pero no suficientes. Por ello, es fundamental llevar a cabo estudios empíricos, tanto de corte cualitativo como cuantitativo para aportar datos a la propuesta de una evaluación de estudiantes capaz de promover una sociedad más justa y equitativa.

8.4. PALABRAS DE CIERRE

Una vez determinadas las lecciones aprendidas de esta investigación, sus fortalezas y debilidades y las líneas futuras que seguir estudiando, nos gustaría terminar esta tesis doctoral con unas reflexiones a modo de aprendizajes y conclusiones finales del trabajo realizado durante estos años.

La realización de este estudio nos permite afirmar que la evaluación es una de las prácticas que más inciden en el desarrollo integral del alumnado, motivo por el cual es necesario seguir investigando en este campo, especialmente por las implicaciones académicas y sociales que presenta.

Los hallazgos presentados en este trabajo ponen de manifiesto que la evaluación marca la dinámica del aula, ya que define la forma de enseñar así como los comportamientos y actitudes de los y las estudiantes. No hay alternativa: si queremos una educación que de verdad contribuya a crear una sociedad más justa, necesitamos una evaluación que trabaje en esa misma dirección.

Seguir profundizando en las concepciones sobre evaluación es imperante. Solamente conociendo las ideas implícitas del profesorado y del alumnado acerca de la práctica evaluativa podremos hacer conscientes a los protagonistas del proceso educativo de su forma de evaluar y así poder transformarla para mejorarla.

En esta línea, el uso de la fenomenografía se consolida como una metodología de investigación poco empleada en nuestro contexto y que nos ayuda a acercarnos a las ideas más personales de los docentes acerca de la evaluación y con eso, entender mejor su práctica profesional. En este sentido, sería

interesante realizar publicaciones metodológicas que diesen a conocer las potencialidades y uso del enfoque fenomenográfico en el campo de la investigación educativa. Asimismo, seguir trabajando en publicar estudios empíricos con resultados acerca de las concepciones sobre evaluación va a promover que sea cada vez una línea de investigación más sólida en España.

Es fundamental abrir un debate entre docentes, administradores, tomadores de decisiones e investigadores sobre qué es una evaluación justa y, con ello, cómo es la evaluación de los estudiantes. Desde nuestro punto de vista, si optamos por un modelo igualitario, con pruebas “objetivas” y sin ser sensible con las diferencias de los alumnos se estará apostando por una educación centrada en la selección y jerarquización de los estudiantes, con lo que se profundiza en las desigualdades sociales.

Nos obstante, si optamos por un modelo basado en la equidad, adaptado a las necesidades de los estudiantes, participativo, con instrumentos variados, cualitativo y esencialmente formativo, se estará apostando por una educación que contribuya a una sociedad más justa. Desde el enfoque de esta tesis doctoral las evidencias son claras: necesitamos que en las escuelas se opte por este tipo de evaluación, una práctica evaluativa que sea equitativa, sensible, crítica y socialmente más justa.

Si y solo si somos utópicos lograremos grandes avances en educación. Es fundamental que tengamos claro que la evaluación es una de las prácticas educativas con más repercusiones a nivel académico, personal y social en la vida de los estudiantes. Solo si tenemos claro el papel político de la evaluación educativa seremos capaces de mirarla con ojos críticos y creer que una evaluación diferente es posible.

Depende de nosotros, como investigadores y como docentes, llevar a cabo una evaluación que empodere al alumnado, que favorezca su desarrollo integral, que sea sensible con sus diferencias, que les haga ser críticos y que promueva que actúen para el cambio social, ese cambio tan necesario en la escuela, en la educación y en el mundo.

*“La utopía es el principio de todo progreso
y el diseño de un futuro mejor”
(Anatole France)*



CAPÍTULO 9.

Conclusions

This chapter presents some conclusive ideas of this research organized in four sections. Firstly, we describe the lessons learned and their implications for a fairer student assessment. Secondly, the following section shows the strengths and positives traits of the study as well as the limitations founded throughout the study development. Thirdly, we track with the possible lines of research that will argument our future work. Finally, some final ideas will close this chapter and all together, it will be this PhD dissertation.

9.1. LESSONS LEARNED OF THE RESEARCH

The previous study has given us a fundamental idea: the relevance of assessment for the integral development of students. What, how, when, where, with whom and for whom of assessment, are decisions that indicate directly and measurable in the students' learning. An inadequate assessment, as we have seen, may even hinder the overall students' development. Hence, it is necessary to put it in the first line of attention in the initial and continuous teacher training.

In this sense, assessment practices affect the most academic performance of students in Primary Education, and thus, it should be further enhanced in the classroom with: a) continuous learning monitoring, b) to review homework and notebooks and give feedback to students, c) to assess students' interest and participation in class, and d) to consider all activities developed during the class. These findings allow us to affirm that if we want that assessment is going to be a useful tool for learning, we must orientate these aspects, including, students in assessment process.

Rethinking assessment involves moving towards more inclusive, participatory and individualized processes in order to know students' progress, moving away from the use of unique quantitative tests and giving way to the use of several assessment techniques eminently qualitative that really allow to know evolution, learning and development of each student.

A second learning of this research is the relevance of the study of assessment conceptions to understand and go in depth on how teachers assess. Investigating their conceptions has an

undoubted value in improving teaching practices. Knowing and understanding how conceptions are and how they are constructed will be possible to modify them (Remesal, 2011).

The way in which teachers conceive their assessment and how it translates into their daily practice are key elements for educational improvement. Being aware of how to assess and what steps are needed to reach a better assessment is one of the challenges of education system, as well as involving more students in the assessment and learning process.

Another of the conclusions of this research is that teachers have different conceptions about a fair assessment. Thus, their conceptions are in a continuum of two poles: equality and equity. While there are teachers who consider that an assessment will be fair only if it is equal and identical for all, others think that evaluative justice refers to adapting assessment to each of the students and their reality and context.

One of the aspects that most affect assessment conceptions (and especially implicit ideas about what a fair assessment is) is the context in which teachers work. It is not unfortunate that current policy decisions that excessively promote external student assessments, even with hard consequences, have a vision of the closest assessment to egalitarian approaches, and this teachers' works in schools situated in favourable contexts. In this way, once again, those students from school in challenging context are the most affect in this process and they have to face with distant assessments of their practice in the classroom.

If we maintain egalitarian assessment approaches we will continue favouring those who have the most and least need. It

is necessary to break with the heterogeneous assessment policies and to convert this assessment practice into an equitable procedure, sensitive with the context of the schools and to the needs and particularities of each student. An equal assessment for all leaves a clear disadvantage to those who, because of their characteristics, need more resources, both personal and material.

It is interesting that context only influence into teachers' conceptions. Conversely, we cannot affirm that relation in students' conceptions. Although, they have conceptions close to equitable or equitable assessment, these conceptions are not determined by the context of school. This brings us the opportunity to reflect how the political assessment pressure on schools, teaching experiences and school culture affects assessment conceptions, clearly marked in the teachers interviewed. If we want teachers that assess in an equitable way, it is fundamental that assessment policies change from being equal for all and attending in a different way to the schools that most need it, avoiding comparisons and *rankings* that not contribute in students' learning and integral development.

Finally, both the previous study and the conceptions' work reflect how much we need to progress in order to develop in classrooms assessments that are truly in line with current approaches about a curriculum that contribute to an education that works for social justice. In spite of speeches, there are still assessment practices characterized by having a hierarchical, authoritarian and controlling view of students and their learning. In this way, it seems urgent to promote professional development of teachers in alternative assessment strategies in both initial and continuous training. Only with an alternative

assessment will be able to have an education that contributes to social change.

9.2. STRENGTHS AND LIMITATIONS OF THE RESEARCH

The development of this PhD dissertation has allowed us to detect certain strengths of the study but also some limitations that are important to highlight.

Firstly, the use of phenomenography in qualitative research is one of the most relevant strengths of this PhD dissertation. The use of this methodology to deepen on assessment conceptions (and especially on fair assessment) has been scarce used in Spanish studies. In this sense, the 73 interviews carried out provide a complete overview of Primary and Secondary Education teachers' and students' conceptions, not deepening in their classroom practices, but also in their implicit ideas about assessment process. Most of the national researches use other quantitative or qualitative methodologies because their aim is to know the teaching work in their classrooms, whereas the interest of our study relies on their conceptions, being able to coincide or not with their assessment practice.

Another strength of this research is the present topic. The research of a fair assessment has been scarce investigated so far from an empirical point of view. Thus, our research provides useful data for knowledge advancement and to improve education. If we want to improve education, it is fundamental to develop fairer assessment processes, being relevant to deepen into protagonists implicit ideas of teaching and learning process.

Closely related to the ideas mentioned above, this work does not focus only on teachers' conceptions, but also on students

ones. The teachers' and students' implicit ideas do not always coincide, providing a vision full of nuances of what should be a fair assessment and what role teachers and pupils plays on it. In this sense, this PhD dissertation has contributed to offer a complementary vision of teachers' and students' conceptions. This aspect has allowed us to deepen in their similarities and differences of the implicit ideas about how assessments practice have to develop to be fair.

Another of the strengths of this research is to give voice to protagonists of educational process. Deepening into the studies focused on assessment conceptions, most of them use a quantitative approach and analysis with Structural Equations Models of their implicit ideas. Our study brings nuances and reflects the ideas of the protagonists about assessment. We are not only interested in locating their assessment conceptions, but also giving presence on their words and understanding their reflections.

The research presented in this PhD dissertation, like any other, is not without limitations. It is important to be aware of strengths, but also of those aspects that could be improved on this study is fundamental if we want to advance in this research area and education process.

The first limitation that we want to emphasize from this PhD dissertation is that the previous study only focused on Primary Education. In this sense, it would be interesting to study the assessment incidence in Secondary Education students and also know which assessment practices determine their academic performance.

In second place, in phenomenographical study, ambiguous, contradictory or changing conceptions of many teachers make

impossible to clearly locate all teachers in one group or another conceptions, and furthermore, this work is incapable to determine what makes teachers have a more egalitarian or equitable fair assessment conceptions. Due to its phenomenographic nature of our study, another of the weaknesses is the exclusive use of a qualitative approach. In this sense, it would be interesting to complement qualitative study with a quantitative approach with a representative national teachers' sample that offered the possibility of generalizing more the results obtained.

In the conceptions study, participants are not the same in the three sub-studies (the first one was focused on the impact of context on assessment conceptions, the second, on the implied ideas of teachers about a fair assessment, and the third, on students' conceptions of a assessment practice). While the first sub-study has ten participants, the second and third sub-studies have 31 teachers and 32 students who participated in phenomenography. It would be good to extend the first sub-study to more teachers' and also extend it to students.

In relation to participants of our research, it would also have been interesting to extend the sample to other Autonomous Communities. For accessibility issues, our research was carried out as a whole with teachers and students from the Community of Madrid, and it was interesting to interview teachers and students from other communities from Spain in order to, on one hand, confirm the findings and compare the results and, on the another hand, know the existence of differences of their implicit ideas of assessment practice.

Despite the weaknesses noted, we believe that the strengths and advances that have been represented in this study provide clear evidence on the importance of our conceptions research.

This is fundamental if we want to improve education by moving towards a fairer assessment in schools.

9.3. FUTURE RESEARCH LINES

Once mentioned lessons learned from this PhD dissertation as well as its weaknesses, we propose some possible research lines that allow us enrich the obtained findings. Undoubtedly, there is still a long way to go into assessment conceptions and, from our perspective, researches could take several directions in order to complete and improve the study presented in this PhD dissertation.

Assuming the strengths and limitations of this PhD dissertation it would be interesting to continue deepening into the most innovative and powerful aspects of this study and to design other lines of research that attenuate or overcome the weaknesses identified.

A possible line of research that may be addressed in future studies would be to continue investigating the impact of the socio-economic context on teachers' conceptions. Conclusions are clear: this context determines conceptions and practices too. Thus, researches shows that school segregation shapes very different teachings depending on the socio-economic context in the face of the supposed undifferentiated treatment by the Administrations of all the centres. Therefore, it would be necessary to carry out a study, both quantitative and qualitative, to determine if the context could influence teachers' assessment practice and their conceptions.

As we indicated in the limitations section, it was missed in this study to complement the phenomenographical data with a quantitative ex post facto study with the same topic. To do

this, it would be interesting to follow the line initiated by Gavin T. L. Brown, adapting and translating Brown's *CoA (III) inventory* (2003, 2004, 2006, 2009) to know both teachers' and students' assessment conceptions. Once inventory was applied, conceptions would be analysed through Structural Equation Models. This analysis would allow knowing what different types of conceptions have both teachers and students. The results of this study could establish relationships with the results obtained in the present phenomenographical study and verify if there is coherence between the quantitative and qualitative approaches. In this sense, the first contacts with New Zealander researcher Gavin T. L. Brown have already begun to initiate with the translation and validation of his inventory in Spanish.

Likewise, it would be interesting to replicate another inventory like Brown's and his colleagues (2003, 2004, 2008, 2009) focused on fair assessment conceptions. Analysing fair assessment conceptions with Structural Equations Models would also contribute to a vision full of nuances that could complement the data contributed in this PhD dissertation through the phenomenography.

Following other studies carried out with great implications for educational practice (Azis, 2014, 2015; James & Pedder, 2006; Prieto, 2008; Vandeyar & Killen, 2007), another relevant study would focus on deepening the relationship between conceptions and assessment practice in the same study. Given that conceptions determine assessment practice would be relevant to know, on the one hand, the teachers' implicit ideas and, on the other hand, the way in which teachers' assess in the classroom through observations. Those studies that have deepened in this research topic determine that there is a gap

between conceptions and teachers' assessment practice, being interesting to do it in the Spanish context and confirm or refute those findings.

Another element that deserves to be studied is to know the incidence of education policies and, in particular, evaluation in assessment conceptions. Administrations determine, to a large extent, assessment activities of schools, being able to influence the implicit ideas of teachers about their assessment practice, and especially, a fair assessment practice. For this reason, it would be essential to continue researching along this line and to provide relevant data highlighting the impact of external tests on the daily assessment practice of teachers. In this sense, it would be interesting to analyse assessment policies of Administration, conducting an *ex post facto* or qualitative study with national and local assessment policy-makers. Once policy-makers' assessment practices and conceptions are known, we could be compared it with those obtained in schools and complement them with a qualitative study to know teachers' implicit ideas about how external assessment influence them to their daily practice.

There is no doubt that it is urgent to continue researching on assessment for social justice. In this sense, as it has been mentioned in previous sections above, we are deepening in theoretical studies about how an assessment has to be to promote a social change, parallel to the study on assessment conceptions. These theoretical advances on assessment for social justice are necessary but not sufficient. For this reason, it is essential to carry out empirical studies, both qualitative and quantitative in order to contribute data to the proposal of an assessment capable of promoting a more just and equitable society.

9.4. CLOSING REMARKS

Once determined lessons learned from this research, its strengths and weaknesses and the future lines that we are going to continue studying, we would like to finish this PhD dissertation with some reflections as learning mode and final conclusions of the work done during these years.

The fulfilment of this study allows us to affirm that assessment is one of the practices that most affect to integral students' development, which is why it is necessary to continue researching in this field, especially for academic and social implications that assessment presents.

The findings introduced in this document show that assessment marks the dynamics of classroom, since it defines the way of teaching as well as students' behaviours and attitudes. There is no alternative: if we want an education that really contributes to create a more just society, we need an assessment that works in the same direction.

Continuing to deepen into assessment conceptions is imperative. Only by knowing teachers' and students' implicit ideas about assessment practice, we could make the protagonists of the educational process aware of how they assess, and thus, we would be able to transform their practice to improve it.

In this line, the use of phenomenography is consolidated as a research methodology scarce used in our context and that helps us to approach more teachers' personal ideas about assessment and with that, to better understand their professional practice. In this sense, it would be interesting to make methodological publications that would reveal the potentialities of the use of

phenomenographical approach in the field of educational research. Also, if we continue publishing empirical studies with results on assessment conceptions, it will help to promote a more solid line of research in Spain.

It is fundamental to open a debate between teachers, administrators, decision-makers and researchers about what a fair assessment is and thus how the students' assessment must be. From our point of view, if we choose an egalitarian model, with "objective" evidence and without being sensitive to students' differences, we will be betting on an education focused on selection and hierarchization of students, creating social inequalities.

However, if we choose an equity-based model, adapted to students' needs, participatory, with varied instruments, qualitative and essentially formative, we will be betting on an education that contributes to a more just society. From the perspective of this PhD dissertation, the evidence is clear: we need schools that choose for this type of assessment, an equitable, sensitive, critical and socially fair assessment practice.

If and only if we are utopian we will achieve great advances in education. It is essential for us to be clear that assessment is one of the educational practices with the greatest impact on the academic, personal and social levels in students' lives. Only if we are clear about the political role of educational assessment will we be able to look at it with critical eyes and believe that a different assessment is possible. It is up to us, as researchers and as teachers, to carry out an assessment that empowers students, fosters their integral development, is sensitive to

their differences, makes them critical and promotes that their actions will be oriented for social change, a change so necessary in school, in education and in the world.

*"Utopia is the beginning of all progress
and the design of a better future"*
(Anatole France)



Referencias

- Adams, M., Bell, L. A. y Griffin, P. (2007). *Teaching for diversity and social justice*. Nueva York, NY: Routledge.
- Aitken, N. (2011). Student voice in fair assessment practice. En C. F. Webber y J. L. Lupart (Eds.), *Leading student assessment* (pp. 175-200). Ámsterdam: Springer.
- Aitken, N. (2012). *Results from the HLC Assessment Academy*. Recuperado de goo.gl/czJGCj
- Alkin, M. C. (1969). *Evaluating the Cost-Effectiveness of Instructional Programs*. Washington, DC: Bureau of Research.
- Álvarez Méndez, J. M. (2001). *Evaluar para conocer, examinar para excluir*. Madrid: Morata.
- Andrade, H. y Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32(2), 159-181.
- Antón Nuño, M. (2012). *Docencia universitaria: concepciones y*

- evaluación de los aprendizajes: estudio de casos*. Tesis doctoral. Universidad de Burgos.
- Applebaum, B. (2004). Social justice education, moral agency, and the subject of resistance. *Educational Theory*, 54(1), 59-72.
- Applebaum, B. (2005). In the name of morality: Moral responsibility, whiteness and social justice education. *Journal of Moral Education*, 34(3), 277-290.
- Atjonen, P. (2014). Teachers' views of their assessment practice. *Curriculum Journal*, 25(2), 238-259.
- Atkinson, T. y Claxton, G. (2000). *The intuitive practitioner: On the value of not always knowing what one is doing*. Buckingham: Taylor & Francis Group.
- Azis, A. (2012). Teachers' conceptions and use of assessment in student learning. *Indonesian Journal of Applied Linguistics*, 2(1), 40-52.
- Azis, A. (2014). *Indonesian junior secondary school teachers' conceptions of assessment: A mixed methods study*. Tesis Doctoral. Victoria Universidad de Wellington, Nueva Zelanda.
- Azis, A. (2015). Conceptions and practices of assessment: a case of teachers representing improvement conception. *Teflin Journal*, 26(2), 129-154.
- Beane, J.A. y Apple, M. W. (1997). La defensa de las escuelas democráticas. En M. W. Apple (Comp.), *Escuelas democráticas* (pp. 145-163). Madrid: Morata.
- Bickel, R. (2007). *Multilevel Analysis for Applied Research: It's Just Regression*. Nueva York, NY: Guilford Press.
- Black, P. y Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation*

- ‡ *Accountability*, 21, 5–31.
- Black, P. y Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Bloom, B. S. (1975). *Evaluation, instruction, and policy making*. París: International Institute for Educational Planning.
- Booth, S. (1997). On phenomenography, learning and teaching. *Higher Education Research & Development*, 16(2), 135-158.
- Boud, D. y Falchikov, N. (2007). Developing assessment for informing judgement. En D. Boud y N. Falchikov (Eds.), *Rethinking Assessment for Higher Education: Learning for the Longer Term* (pp. 181-197). Londres: Routledge.
- Bowden, J. A. y Walsh, E. (2000). Phenomenography. *Phenomenography*, 1, 154-169.
- Bowden, J., D'All'Alba, G., Martin, E., Masters, G., Laurillard, D., Marton, F., Ramsden, P. y Stephanou, A. (1992). Displacement, velocity and frames of reference: Phenomenographic studies of students' understanding and some implications for teaching and assessment. *American Journal of Physics*, 60, 262-268.
- Bowden, J. (1996). Phenomenographic research - Some methodological issues. En G. Dall'Alba y B. Hasselgren (Eds.), *Reflections on Phenomenography: towards a methodology* (pp. 49-66). Goteburgo: Acta Universitatis Gothoburgensis.
- Bowden, J. A. y Walsh, E. (2000). Phenomenography. *Phenomenography*, 1, 154-169.
- Bransford, D., Brown, A. y Cocking, R. (1999). *How People Learn. Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.

- Brown, G. T. L. (2003). *Teachers' conceptions of assessment*. Tesis Doctoral. Universidad de Auckland, Nueva Zelanda.
- Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301-318.
- Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged version 1, 2. *Psychological Reports*, 99(1), 166-170.
- Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. Auckland: Nova Science Publishers.
- Brown, G. T. L. (2011). Self-regulation of assessment beliefs and attitudes: A review of the Students' Conceptions of Assessment inventory. *Educational Psychology*, 31(6), 731-748.
- Brown, G. T. L. y Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education*, 2, 1-19.
- Brown, G. T. L. y Harris, L. (2012). Student conceptions of assessment by level of schooling: further evidence for ecological rationality in belief systems. *Australian Journal of Educational & Developmental Psychology*, 12, 46-59.
- Brown, G. T. L. y Hirschfeld, G. H. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3-17.
- Brown, G. T. L. y Michaelides, M. (2011). Ecological rationality in teachers' conceptions of assessment across samples from Cyprus and New Zealand. *European Journal of Psychological Education*, 26(3), 319-337.

- Brown, G. T. L. y Lake, R. (noviembre, 2006). Queensland teachers' conceptions of teaching, learning, curriculum and assessment: Comparisons with New Zealand teachers. Comunicación presentada en la *Australian Association for Research in Education Conference*, Adelaide, Australia.
- Brown, G. T. L. y Remesal, A. (2012). Prospective teachers' conceptions of assessment: A cross-cultural comparison. *The Spanish Journal of Psychology*, 15(1), 75-89.
- Brown, G. T. L., Lake, R. y Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education*, 27(1), 210-220.
- Brown, G. T. L., Pishghadam, R. y Sadafian, S. S. (2014). Iranian university students' conceptions of assessment: Using assessment to self-improve. *Assessment Matters*, 6, 5-33.
- Brown, G. T. L., Irving, S. E., Peterson, E. R. y Hirschfeld, G. H. (2009). Use of interactive-informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning and Instruction*, 19(2), 97-111.
- Brown, G. T. L., Hui, S. K., Flora, W. M. y Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5), 307-320.
- Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S. y Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347-363.

- Carlisle, L. R., Jackson, B. W. y Alison, G. (2006). Principles of social justice education: The social justice education in schools project. *Equity & Excellence in Education*, 39(1), 55-64.
- Carlson, C. D. y Francis, D. J. (2002). Increasing the reading achievement of at-risk children through direct instruction: evaluation of the Rodeo Institute for Teacher Excellence. *Journal of Education for Students Placed at Risk*, 7(2), 141-166.
- Casanova, M. A. (1998). Evaluación: Concepto, tipología y objetivos. En M. A. Casanova (Ed.), *La evaluación educativa* (pp. 67-102). Ciudad de México: Biblioteca para la Actualización del Maestro.
- Chadwick, C. y Rivera, N. (1990). El concepto de evaluación y evaluación formativa. En C. Chadwick, C. y N. Rivera, (Coords.), *Evaluación formativa para el docente* (pp. 33-62). Barcelona: Paidós.
- Chen, J. A., Morris, D. B. y Mansour, N. (2015). Science teachers' beliefs: Perceptions of efficacy and the nature of scientific knowledge and knowing. En H. Fives y M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 370-386). Londres: Routledge.
- Chilisa, B. (2000). Towards equity in assessment: Crafting gender-fair assessment. *Assessment in Education: Principles, Policy & Practice*, 7(1), 61-81.
- Cline, T. (2000). *Principles and Practise of Fair Assessment*. Eric Digest.
- Cochran-Smith, M., Shakman, K., Jong, C., Terrell, D. G., Barnatt, J. y McQuillan, P. (2009). Good and just teaching: The case for social justice in teacher education. *American Journal of Education*, 115(3), 347-377.

- Cohen, L., Manion, L. y Morrison, K. (2000). *Research Methods in Education*. Londres: Routledge Falmer.
- Coll, C. (1999). *Psicología de la instrucción: la enseñanza y el aprendizaje en la Educación Secundaria*. Barcelona: ICE Universitat Barcelona.
- Coll, C. y Remesal, A. (2009). Mathematics teachers' conceptions about the functions of assessment in compulsory education. *Journal for the Study of Education and Development*, 32(3), 391-404.
- Connell, R. W. (1997). *Escuelas y Justicia Social*. Madrid: Morata.
- Corbin, J. y Strauss, A. (2008). *Basics of qualitative research. Technics and procedures for developing grounded theory*. Londres: Sage Publications.
- Cronbach, L. (1963). Course improvements through evaluation. *The Teachers College Record*, 64(8), 672-672.
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess., R. D., Hornik., R. C., Phillips, D. C., Walker, D. F. y Weiner, S. S. (1980). *Toward reform of program evaluation*. San Francisco, CA: Jossey-Bass.
- Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review*, 64(1), 5-31.
- Davidson, J. (2004). *Evaluation methodology basics. The nuts and bolts of sound evaluation*. Londres: Sage.
- Davis, D. S. y Neitzel, C. (2010). The relationship between students' reading orientations and their strategic activity during a collaborative reading task. *Reading Psychology*, 31(6), 546-579.
- Delandshere, G. (2010). Implicit theories, unexamined

- assumptions and the status quo of educational assessment. *Assessment in Education: Principles, Policy y Practice*, 8(2), 113-133.
- Descy, P. y Tessaring, M. (2004). *Impact of education training*. París: Centre Européen pour le Developpement de la Formation Professionnelle.
- Dortins, E. (2002). Reflections on phenomenographic process: Interview, transcription and analysis. *Quality Conversations: Research and Development in Higher Education*, 25, 207-213.
- Doval, L. (1995). Aportaciones del diagnóstico en educación. De la historiografía al concepto. *Revista de Investigación Educativa*, 13(26), 201-216.
- Dubet, F. (2011). *Repensar la Justicia Social*. Madrid: Siglo XXI editores.
- Eisner, E. W. (1985). *The art of educational evaluation. A personal view*. Londres: The Falmer Press.
- Ekeblad, E. y Bond, C. (noviembre, 1994). The nature of a conception: Questions of context. Comunicaciçon presentada en el congreso *Phenomenography: Philosophy and Practice, Proceedings*, Queensland, Australia.
- Escudero, T. (2003). Desde los tests hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo de la evaluación en educación. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 9(1), 11-43.
- Fives, H. y Gill, M. G. (2014). *International handbook of research on teachers' beliefs*. Nueva York, NY: Routledge.
- Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P. y Rees, M. (2012). Faculty and students conceptions of

- assessment in higher education. *Higher Education*, 64(1), 119-133.
- Fraser, N. (2008). La justicia social en la era de la política de identidad: redistribución, reconocimiento y participación. *Revista de Trabajo*, 4(6), 83-99.
- Freire, P. (1970). *Pedagogía del Oprimido*. Ciudad de México: Editorial Paz y Tierra.
- García Garduño, J. M. (2005). El avance de la evaluación en México y sus antecedentes. *Revista Mexicana de Investigación Educativa*, 10(27), 1275-1283.
- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, 18, 694-734.
- Gipps, C. y Stobart, G. (2009). Fairness in assessment. En C. Wyatt-Smith y J. Cumming (Eds.), *Educational assessment in the 21st century* (pp. 105-118). Ámsterdam: Springer.
- Giroux, H. A. (2001). *Cultura, política y práctica educativa*. Barcelona: Graó.
- Giroux, H. A. (1990). *Los profesores como intelectuales*. Barcelona: Paidós.
- Giroux, H. A. (1992). *Teoría y resistencia en educación: una pedagogía para la oposición*. Madrid: Siglo XXI Editores.
- González-Ugalde, C. (2014). Investigación fenomenográfica. Magis. *Revista Internacional de Investigación en Educación*, 7(14), 141-158.
- Glaser, B. S. y Strauss, A. (1971). *The discovery of grounded theory*. Nueva York, NY: Transaction Publishers.
- Goldstein, H. (1995). *Multilevel statistical models*. Londres: Edwards Arnold.

- Goodwin, A. L. (2012). *Assessment for equity and inclusion: Embracing all our children*. Londres: Routledge.
- Greene, J. C. (2006). Evaluation, democracy, and social change. En I. Shaw, J. Greene y M. Mark (Eds.), *The Sage handbook of evaluation* (pp. 118-140). Nueva York, NY: Sage.
- Griffiths, M. (2003). *Action for Social Justice in Education*. Buckingham: Open University Press.
- Griffiths, T., Gore, J., y Ladwig, J. (noviembre, 2006). Teachers' fundamental beliefs, commitment to reform, and the quality of pedagogy. Comunicación presentada en la Conferencia anual *Proceedings Australian Association for Research in Education*, Adelaide, Australia.
- Gronlund, N. E. (1985). *Stating objectives for classroom instruction*. Nueva York, NY: Macmillan Publishing Company.
- Gronlund, N. E. (2003). *Assessment of Student Achievement*. Needham Heights, MA: Allyn & Bacon Publishing.
- Guba, E. G. y Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.
- Hackman, H. W. (2005). Five essential components for social justice education. *Equity & Excellence in Education*, 38(2), 103-109.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning-tensions and synergies. *Curriculum Journal*, 16(2), 207-223.
- Harris, L. R. (2008). A phenomenographic investigation of teacher conceptions of student engagement in learning. *The Australian Educational Researcher*, 35(1), 57-79.

- Harris, L. y Brown, G. T. L. (octubre, 2008). *New Zealand teachers' conceptions of the purpose of assessment: Phenomenographic analyses of teachers' thinking*. Comunicación presentada en la *Australian Association for Research in Education Conference*, Brisbane, Australia.
- Harris, L. y Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practices*, 16(3), 365-381.
- Harris, L. R., Brown, G. T. L. y Harnett, J. A. (2014). Understanding classroom feedback practices: A study of New Zealand student experiences, perceptions, and emotional responses. *Educational Assessment, Evaluation and Accountability*, 26(2), 107-133.
- Hasselgren, B. y Beach, D. (1997). Phenomenography-a "good-for-nothing brother" of phenomenology? Outline of an analysis. *Higher Education Research & Development*, 16(2), 191-202.
- Hernández-Castilla, R. y Salinas, B. (2008). La evaluación didáctica como componente del diseño/desarrollo curricular. En A. Herrán y J. Paredes (Coords.), *Didáctica General* (pp. 235-248). Madrid: McGrawhill.
- Hidalgo, N. y Murillo, F. J. (2015). Enfoques alternativos de evaluación de los estudiantes. En J. Domingo y M. Pérez Ferra (Coords.), *Aprendiendo a enseñar. Manual Práctico de Didáctica* (pp. 171-174). Madrid: Pirámide.
- Hidalgo, N. y Murillo, F. J. (2016). Evaluación de estudiantes para la Justicia Social. Propuesta de un Modelo. *Revista Internacional de Educación para la Justicia Social*, 5(2), 159-179.

- Hidalgo, N. y Murillo, F. J. (2017). Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 15(1), 107-128.
- Hidri, S. (2015). Conceptions of assessment: Investigating what assessment means to secondary and university teachers. *Arab Journal of Applied Linguistics*, 1(1), 19-43.
- Hirschfeld, G. H. y Brown, G. T. L. (2009). Students' conceptions of assessment: Factorial and structural invariance of the SCoA across sex, age, and ethnicity. *European Journal of Psychological Assessment*, 25(1), 30-38.
- Honneth, A. (2003). Redistribution as recognition: A response to Nancy Fraser. En N. Fraser y A. Honneth (Eds.), *Redistribution or recognition: A political-philosophical debate* (pp. 110-197). Londres: Verso.
- Honneth, A. (2007). *Reificación: un estudio en la teoría del reconocimiento*. Madrid: Katz Editores.
- Hox, J. J. (1995). *Applied multilevel Analysis*. Amsterdam: TT-Publikaties.
- Huba, M. E. y Freed, J. E. (2000). Learner centered assessment on college campuses: Shifting the focus from teaching to learning. *Community College Journal of Research and Practice*, 24(9), 759-766.
- Hunt, G. H., Wiseman, D. G. y Touzel, T. J. (2009). *Effective teaching: preparation and implementation*. Springfield, IL: Charles C. Thomas.
- ICAP. (1995). *What Is Evaluation?*. Recuperado de <http://www.icap.org/PolicyTools/Toolkits/EvaluationToolkit/2WhatIsEvaluation/tabid/441/Default.aspx>

- James, M. y Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109-138.
- Janesick, V. J. (2008). Reflexiones sobre la violencia de las pruebas de evaluación estandarizada y la naturaleza tranquilizadora de la pedagogía crítica. En P. McLaren y J. L. Kincheloe (Eds.), *Pedagogía Crítica. De qué hablamos, dónde estamos* (pp. 323- 337). Barcelona: Graó.
- Johansson, B., Marton, F. y Svensson, L. (1985). An approach to describing learning as change between qualitatively different conceptions. En A. L. Pines y L. H. T. West (Eds.), *Cognitive structure and conceptual change* (pp. 45-87). Nueva York, NY: Academic Press.
- Jornet, M. J. (2016). Dimensiones docentes y cohesión social: reflexiones desde la evaluación. *Revista Iberoamericana de Evaluación Educativa*, 5(1e), 350-362.
- Jornet, J. M., Sancho-Álvarez, C. y Bakieva, M. (2015). Diseño y validación del constructo teórico de Justicia Social percibida por el alumnado en los centros escolares. *Revista Internacional de Educación para la Justicia Social*, 4(2), 111-126.
- Killen, R. (2005). Programming and assessment for quality teaching and learning. Sidney: Thomsom.
- Kohl, H. (2000). Teaching for Social Justice. *Rethinking Schools*, 15(2), 14-35.
- Kumashiro, K. K. (2015). *Against common sense: Teaching and learning toward social justice*. Londres: Routledge.
- Kvale, S. (1996). *Inter Views. An Introduction to Qualitative Research Interviewing*. Nueva York, NY: Sage Publications, London.

- Kyriakides, L., Creemers, B. P. M. y Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25, 12-23.
- LaCueva, A. (1997). La evaluación en la escuela: una ayuda para seguir aprendiendo. *Revista da Faculdade de Educação*, 23(1-2).
- Lam, T. (1995). *Fairness in Performance Assessment*. Recuperado de <http://www.ericdigests.org/1996-4/fairness.htm>
- Landy, F. J., Barnes-Farrell, J. L. y Cleveland, J. N. (1980). Perceived fairness and accuracy of performance evaluation: A follow up. *Journal of Applied Psychology*, 63, 355-356.
- Lukas, J. F. y Santiago, K. (2004). *Evaluación educativa*. Madrid: Alianza Editorial.
- Ma, X., Ma, L. y Bradley, K. D. (2008). Using multilevel modeling to investigate school effects. En A. A. O'connell y B. McCoach (Eds.), *Multilevel Modeling of Educational Data* (pp. 59-110). Nueva York, NY: Information Age Publishing Inc.
- Madaus, G. (1991). The effects of important tests on students: Implications for a national examination system. *The Phi Delta Kappan*, 73(3), 226-231.
- Mahncke, M. (2010). *Enfoques de aprendizaje y de estudio de los estudiantes universitarios*. Tesis doctoral. Universidad Ramón Llull, Barcelona.
- Marshall, B. y Drummond, M. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133-149.

- Martín, E., Mateos, M., Martínez, P., Cervi, J., Pecharromán, R. y Villalón, R. (2006). Las concepciones de los profesores de educación primaria sobre la enseñanza y el aprendizaje. En J. I. Pozo, N. Scheuer, M. Pérez Echeverría, M. Mateos, E. Martín y M. De la Cruz (Eds.), *Nuevas formas de pensar la enseñanza y el aprendizaje. Las concepciones de profesores y alumnos* (pp. 171-187). Barcelona: Graó.
- Martínez-Garrido, C. y Murillo, F. J. (2013). El uso de los Modelos Multinivel en la Investigación Educativa. Estadísticas avanzadas para conocer y cambiar la educación en América Latina. En A. Salcedo (Ed.), *Estadística en la Investigación: competencia transversal en la formación universitaria* (pp. 47-70). Caracas: UCV.
- Martínez-Garrido, C. y Murillo, F. J. (2014). Programas para la realización de Modelos Multinivel. Un análisis comparativo entre MLwiN, HLM, SPSS y Stata. *REMA. Revista Electrónica de Metodología Aplicada*, 9(2), 1-24.
- Marton, F. (1981). Phenomenography-Describing conceptions of the world around us. *Instructional Science*, 10, 177-200.
- Marton, F. (1986). Phenomenography: A research approach to investigating different understandings of reality. *Journal of Thought*, 21(3), 28-49.
- Marton, F. (1996). Cognosco ergo sum. Reflections on reflections. *Reflections on phenomenography: Toward a Methodology*, 109, 163-89.
- McLaren, P. (1984). *La vida en las escuelas: una introducción a la pedagogía crítica en los fundamentos de la educación*. Ciudad de México: Siglo XXI editores.
- McLaren, P. (1995). *Pedagogía crítica y cultura depredadora*. Madrid: Paidós.

- McLaren, P. (2010). Revolutionary critical pedagogy. *InterActions: UCLA Journal of Education and Information Studies*, 6(2), 2-11.
- McMillan, J. H. y Schumacher, S. (2006). *Research in Education: Evidence-Based Inquiry*. Boston, MA: Pearson Education.
- Mertens, D. M. (2007). Transformative Paradigm Mixed Methods and Social Justice. *Journal of Mixed Methods Research*, 1(3), 212- 225.
- Moreno Olivos, T. (2011). La cultura de la evaluación y la mejora de la escuela. *Perfiles educativos*, 33(131), 116-130.
- Moreno Olivos, T. (2014). Posturas epistemológicas frente a la evaluación y sus implicaciones en el currículum. *Perspectiva Educacional*, 53(1), 3-18.
- Moyano-Díaz, E. (2016). Construyendo el concepto de felicidad para adultos desde la fenomenografía. *Universum (Talca)*, 31(2), 141-156.
- Muijs, D., Chapman, C., Collins, A. y Armstrog, P. (2010). *Maximum Impact Evaluation*. Manchester: University of Manchester.
- Murillo, F. J. (2006). *Evaluación del desempeño y carrera profesional docente. Una panorámica de América y Europa*. Santiago de Chile: UNESCO.
- Murillo, F. J. (2007a). School Effectiveness Research in Latin America. En T. Townsend (Ed.), *International Handbook of School Effectiveness and Improvement* (pp. 75-92). Nueva York, NY: Springer.
- Murillo, F. J. (2007b). Investigación iberoamericana sobre *eficacia escolar*. Bogotá: Convenio Andrés Bello.

- Murillo, F. J. (2008). Los modelos multinivel como herramienta para la investigación educativa. *Magis, Revista Internacional de Investigación en Educación*, 1, 45-62.
- Murillo, F. J. (2010). ¿Quiénes son los Responsables de los Resultados de las Evaluaciones?: Hacia un Planteamiento de Valor Agregado en Educación. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 8(4), 3-9.
- Murillo, F. J. (2016). De la docencia eficaz a la evaluación eficaz de la docencia. En G. Guevara, M. T. Meléndez, F. E. Ramón, H. Sánchez y F. Tirado (Coords.), *La evaluación docente en el mundo* (pp. 200-225). Ciudad de México: Fondo de Cultura Económica.
- Murillo, F. J. y Hernández-Castilla, R. (2011). Hacia un concepto de Justicia Social. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 9(4), 7-23.
- Murillo, F. J. y Hernández-Castilla, R. (2014). Liderando Escuelas Justas para la Justicia Social.. *Revista Internacional de Educación para la Justicia Social*, 3(2), 13-32.
- Murillo, F. J. e Hidalgo, N. (2014). Evaluación justa: entre la igualdad y la equidad. *Revista Iberoamericana de Evaluación Educativa*, 7(3), 5-6.
- Murillo, F. J. e Hidalgo, N. (2015a). Dime cómo evalúas y te diré qué sociedad construyes. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 5-9.
- Murillo, F. J. e Hidalgo, N. (2015b). La evaluación de estudiantes como práctica política y social en tiempos neoliberales. *Revista Iberoamericana de Evaluación Educativa*, 8(2), 5-7.

- Murillo, F. J. e Hidalgo, N. (2015c). Enfoques fundamentantes de la Evaluación de Estudiantes para la Justicia Social. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 43-61.
- Murillo, F. J. e Hidalgo (2016b). Hacia un proceso de evaluación docente justo y socialmente justo. *Revista Iberoamericana de Evaluación Educativa* 9(2), 5-8.
- Murillo, F. J. e Hidalgo, N. (2017). Students' conceptions about a fair assessment of their learning. *Studies in Educational Evaluation*, 53, 10-16.
- Murillo, F. J. e Hidalgo, N. (en prensa). Las concepciones sobre una evaluación justa de los estudiantes. Un estudio fenomenográfico desde la perspectiva de los docentes. *Revista Complutense de Educación*, en prensa.
- Murillo, F. J. y Román, M. (2008). Resultados de aprendizaje en América Latina a partir de las evaluaciones nacionales. *Revista Iberoamericana de Evaluación Educativa* 1(1), 6-35.
- Murillo, F. J. y Román, M. (2009). Mejorar el desempeño de los estudiantes de América Latina: algunas reflexiones a partir de los resultados del SERCE. *Revista Mexicana de Investigación Educativa*, XIV(41), 451-484.
- Murillo, F. J. y Román, M. (2010). Retos en la evaluación de la calidad de la educación en América Latina. *Revista Iberoamericana de Educación*, 53, 97-120.
- Murillo, F. J., Hidalgo, N. y Flores, S. (2016). Incidencia del contexto socio-económico en las concepciones docentes sobre evaluación. *Profesorado. Revista de Currículum y Formación del Profesorado*, 20(3), 251-281.
- Murillo, F. J., Martínez-Garrido, C. y Hernández-Castilla, R. (2011). Decálogo para una enseñanza eficaz. *REICE*.

- Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 9(1), 6-27.
- Murillo, F. J., Martínez-Garrido, C. e Hidalgo N. (2014). Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España. *Estudios sobre Educación*, 27, 91-113.
- Murillo, F. J., Román, M. y Hernández-Castilla, R. (2011). Evaluación educativa para la Justicia Social. *Revista Iberoamericana de Evaluación Educativa*, 4(1), 7-23.
- Murillo, F.J., Hernández-Castilla, R., Hidalgo, N. y Martínez-Garrido, C. (2014). Elaboración y Evaluación Psicométrica de la Escala de Actitudes hacia la Justicia Social en Educación (EAJSE). *Revista Internacional de Educación para la Justicia Social*, 3(2), 215-233.
- Nevo, D. (1983). The conceptualization of educational evaluation: An analytical review of the literature. *Review of Educational Research*, 53(1), 117-128.
- North, C. E. (2006). More than words? Delving into the substantive meaning (s) of “social justice” in education. *Review of Educational Research*, 76(4), 507-535.
- Olmos, S. (2008). Evaluación formativa y sumativa de estudiantes universitarios: aplicación de las tecnologías a la evaluación educativa. *Revista de Docencia Universitaria*, 9(1), 159-173.
- Opre, D. (2010). *Expertiza didáctica universitară*. Praga: ASCR.
- Opre, D. (2015). Teachers’ conceptions of assessment. *Procedia - Social and Behavioral Sciences* 209, 229-233.
- Orgill, M. (2012). Phenomenography. En N. M. Seed (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 2608-2611). Nueva York, NY: Springer.

- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332.
- Patton, M. Q. (2003). Utilization-focused evaluation. En T. Kellaghan, D. L. Stufflebeam y L. A. Wingate (Eds.). *International Handbook of Educational Evaluation* (pp. 223-242). Dordrecht: Kluwer.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.
- Pellegrino, J. W., Chudowsky, N. y Glaser, R. (2001). Knowing what students know. *Issues in Science and Technology*, 19(2), 48-52.
- Pepper, M. B. y Pathak, S. (2008). Classroom contribution: What do students perceive as fair assessment?. *Journal of Education for Business*, 83(6), 360-368.
- Pérez Expósito, L. y González Aguilar, D. A. (2011). "Dime cómo evalúas y te diré qué enseñas". Un análisis teórico sobre las relaciones entre la evaluación del aprendizaje y la enseñanza-aprendizaje de la justicia Social. *Revista Iberoamericana de Evaluación Educativa*, 4(1), 135-148.
- Peterson, E. R. e Irving, S. E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction*, 18(3), 238-250.
- Petrou, A., Angelides, P. y Leigh, J. (2009). Beyond the difference: From the margins to inclusion. *International Journal of Inclusive Education*, 13(5), 439-448.
- Pettifor, J. L. y Saklofske, D. H. (2012). Fair and ethical student assessment practices. En C. F. Webber y J. L. Lupart (Eds.), *Leading student assessment* (pp. 87-106). Amsterdam: Springer.

- Philipp, R. A. (2007). Mathematics Teachers. Beliefs and A ect'. En F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 257-315). Charlotte, NC: Information Age Publishing.
- Picón Jácome, É. (2013). La rúbrica y la justicia en la evaluación. *Íkala, Revista de Lenguaje y Cultura*, 18(3), 79-94.
- Pope, N., Green, S. K., Johnson, R. L. y Mitchell, M. (2009). Examining teacher ethical dilemmas in classroom assessment. *Teaching and Teacher Education*, 25(5), 778-782.
- Popham, W. J. (1980). *Problemas y técnicas de la evaluación educativa*. Madrid: Anaya.
- Popham, W. J. (1990). *Modern educational measurement: A practitioner's perspective*. Boston, MA: Allyn and Bacon.
- Popham, W. J. (2004). Why Assessment Illiteracy Is Professional Suicide. *Educational Leadership*, 62(1), 82.
- Pozo, J. (2006). La cultura del aprendizaje en la sociedad del conocimiento. En J. Pozo, N. Scheuer, M. P. Pérez Echevarría, M. Mateos, E. Martín y M. De la Cruz (Eds.), *Nuevas formas de pensar la enseñanza y el aprendizaje. Las concepciones de profesores y alumnos* (pp. 29-53). Barcelona: Graó.
- Pozo, J., Scheuer, N., Mateos, M. y Pérez Echeverría, M. P. (2006). Las teorías implícitas sobre la enseñanza el aprendizaje. En J. Pozo, N. Scheuer, M. P. Pérez Echevarría, M. Mateos, E. Martín y M. De la Cruz (Eds.), *Nuevas formas de pensar la enseñanza y el aprendizaje. Las concepciones de profesores y alumnos* (pp. 95-134). Barcelona: Graó.

- Pratt, D. D. (1992). Conceptions of teaching. *Adult Education Quarterly*, 42(4), 203-220.
- Prieto, M. (2008). Creencias de los profesores sobre evaluación y efectos incidentales. *Revista de Pedagogía*, 29(84), 123-144.
- Prieto, M. y Contreras, G. (2008). Las concepciones que orientan las prácticas evaluativas de los profesores: un problema a develar. *Estudios Pedagógicos (Valdivia)*, 34(2), 245-262.
- Quesada-Serra, V., Rodríguez-Gómez, G. e Ibarra-Sáiz, M. S. (2016). What are we missing? Spanish lecturers' perceptions of their assessment practices. *Innovations in Education and Teaching International*, 53(1), 48-59.
- Rawls, J. (1971). *Teoría de la Justicia*. Ciudad de México: FCE.
- Rawls, J. (2002). *La justicia como equidad. Una reformulación*. Barcelona: Paidós.
- Reed, B. I. (junio, 2006). Phenomenography as a way to research the understanding by students of technical concepts. Comunicación presentada en el *International Seminar of Technological Innovation Sustainability*. Barcelona, España.
- Reise, S. P. y Duan, N. (2003). *Multilevel Modeling: Methodological Advances, Issues, and Applications*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Remesal, A. (2006). Los problemas en la evaluación del aprendizaje matemático en la educación obligatoria: perspectiva de profesores y alumnos. Tesis doctoral. Universidad de Barcelona, España.

- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472-482.
- Richardson, J. T. (1999). The concepts and methods of phenomenographic research. *Review of Educational Research*, 69(1), 53-82.
- Rodríguez-Gómez, G. e Ibarra-Sáiz, M.S. (2012). La evaluación de los procesos y contextos educativos. En F. Navaridas Nalda (Ed.), *Procesos y contextos educativos* (pp. 257-298). Murcia: Genuve Ediciones.
- Rodríguez-Gómez, G. e Ibarra-Sáiz, M. S. (2015). Assessment as Learning and empowerment: towards sustainable learning in higher education. En M. Peris-Ortiz, J. M-Merigo Lindahl (Eds.), *Sustainable Learning in Higher Education* (pp. 1-20). Nueva York, NY: Springer International Publishing.
- Rodríguez-Gómez, G., Ibarra-Sáiz, M. S., Gallego-Noche, B., Gómez-Ruiz, M. Á. y Quesada-Serra, V. (2012). La voz del estudiante en la evaluación del aprendizaje: un camino por recorrer en la universidad. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 18(2), 133-153.
- Rodríguez-Gómez, G., Ibarra-Sáiz, M. S. e Ibáñez, J. (en prensa). Competencias básicas en evaluación. Un estudio sobre la percepción de los estudiantes universitarios. *Educación XXI*.
- Rogers, T. (2010). Principles for Fair Student Assessment Practices for Education in Canada. *Assessment in Education*, 3(3), 413-432.
- Román, M. y Murillo, F. J. (2009). Learning assessment in Latin America: School performance behavior and trends

- of Latin American pupils in primary and secondary education. *Sísifo, Educational Sciences Journal*, 9, 31-46.
- Rossi, P. H., Lipsey, M. W. y Freeman, H. E. (2004). *Evaluation: A systematic approach*. Thousand Oaks, CA: Sage Publications.
- Rubie-Davies, C. (2015). Teachers' instructional beliefs and the classroom climate. En H. Fives y M.G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 266-283). Nueva York, NY: Routledge.
- Rubie-Davies, C. M., Flint, A. y McDonald, L. G. (2012). Teacher beliefs, teacher characteristics, and school contextual factors: What are the relationships?. *British Journal of Educational Psychology*, 82(2), 270-288.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, 5(1), 77-84.
- Säljö, R. (1996). Minding action - Conceiving of the world versus participating in cultural practices. En G. Dall'Alba y B. Hasselgren (Eds.), *Reflections on phenomenography: toward a methodology?* (pp. 19-33). Goteburgo: Goteborg Studies in Educational Sciences.
- Säljö, R. (1994). Minding action: Conceiving of the world versus participating in cultural practices. *Nordisk Pedagogik*, 2, 71-80.
- Sambell, K., McDowell, L. y Brown, B. (1997). 'But is it fair?': An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, 23(4), 349-371.
- Sancho-Álvarez, C., Jornet, J. M. y González-Such, J. (2016). El constructo Valor Social Subjetivo de la Educación: validación cruzada entre profesorado de escuela y

- universidad. *Revista de Investigación Educativa*, 34(2), 329-350.
- Sancho-Álvarez, C., Jornet, J. M. y Perales Montolío, M. J. (2014). Valor social subjetivo de la educación: docencia y resultados escolares. *Revista Iberoamericana de Evaluación Educativa*, 7(2), 197-207.
- Sandbergh, J. (1997). Are phenomenographic results reliable?. *Higher Education Research & Development*, 16(2), 203-212.
- Sanmartí, N. (2007). *Evaluar para aprender*. Barcelona: Graó.
- Santos Guerra, M. Á. (2003). *Una flecha en la diana: la evaluación como aprendizaje*. Madrid: Narcea.
- Schraw, G. y Olafson, L. (2003). Teachers' epistemological world views and educational practices. *Journal of Cognitive Education and Psychology*, 3(2), 178-235.
- Scott, S., Webber, C., Lupart, J., Aitken, N. y Scott, D. (2014). Fair and equitable assessment practices for all students. *Assessment in Education: Principles, Policy & Practice*, 21(1), 52-70.
- Scriven, M. (1991). Beyond formative and summative evaluation. En M. W. McLaughlin y D. C. Phillips (Eds.), *Evaluation and education at quarter century* (pp. 19-64). Chicago, IL: National Society for the Study of Education.
- Segers, M. y Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment?. *Studies in Educational Evaluation*, 37(1), 49-54.
- Sensoy, Ö. y DiAngelo, R. (2014). Respect differences? Challenging the common guidelines in social justice education. *Democracy and Education*, 22(2), 1-10.
- Sephard, L. A. (2006). Assessment in the classroom. En R. L.

- Brennan (Ed.), *Educational Measurement*, (pp. 623-646). Ciudad de México: ACE.
- Sephard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Sjöström, B. y Dahlgren, L. O. (2002). Applying phenomenography in nursing research. *Journal of Advanced Nursing*, 40(3), 339-345.
- Stake, R. (1967). The countenance of educational evaluation. *The Teachers College Record*, 68(7), 523-540.
- Stevenson, H. P. (2007). A case study in leading schools for social justice: when morals and markets collide. *Journal of Educational Administration*, 45(6), 769-781.
- Stiggins, R. J. (2001). *Student-involved classroom assessment*. Nueva York, NY: Prentice Hall.
- Stiggins, R. y DuFour, R. (2009). Maximizing the power of formative assessment. *Phi Delta Kappan*, 90(9), 640-644.
- Stobart, G. (2005). Fairness in multicultural assessment systems. *Assessment in Education*, 12(3), 275-287.
- Stufflebeam D. L. (1971). *Educational evaluation and decision making*. Ithaca, IL: Peacock.
- Stufflebeam, D. L. y Shinkfield, A. (1987). *Evaluación sistemática*. Barcelona: Paidós.
- Stufflebeam, D. L. y Shinkfield, A. (1995). *Evaluación sistemática: Guía teórica y práctica*. Madrid: Ediciones Paidós Ibérica.
- Susinos, T. (2012). Las posibilidades de la voz del alumnado para el cambio y la mejora educativa. *Revista de Educación*, 359(3), 16-23.
- Suskie, L. (2002). Fair assessment practices: Giving students equitable opportunities to demonstrate learning.

Adventures in Assessment, 14, 5–10.

- Sutton, R. (2009). Formative Assessment the Way the School Does Business: The impact and Implications of formative assessment for teachers, students and school leaders. En A. Hargreaves, A. Lieberman, M. Fullan y D. Hopkins (Eds.), *Second International Handbook of Educational Change* (pp. 883-899). Londres: Springer.
- Syverson, M. A. (2009). Social Justice and Evidence-Based Assessment with the Learning Record. *Forum on Public Policy*, 42(2), 109-117.
- Taba, H. (1962). *Curriculum development. Theory and practice*. Nueva York, NY: Harcourt Brace.
- Taras, M. (2005). Assessment—summative and formative—some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466-478.
- Taylor, C. (2003). *El multiculturalismo y “la política del reconocimiento”*. Ciudad de México: Fondo de Cultura Económica.
- Taylor, S. J. y Bogdan, R. (1996). La entrevista en profundidad. En S. J. Taylor y R. Bogdan (Eds.), *Introducción a los métodos cualitativos de investigación* (pp. 100-132). Madrid: Nova-Gráfica.
- Teddlie, C. y Reynolds, D. (2000). *The International Handbook on School Effectiveness Research*. Nueva York, NY: Routledge Falmer
- TenBrink, T. (1974). *Evaluation: a Practical Guide for Teachers*. Nueva York, NY: McGraw Hill.
- Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. En D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning*

- (pp. 127-146). Nueva York, NY: Macmillan.
- Tierney, R. D. (2013). Fairness in classroom assessment. En J. Mcmillan (Ed.), *Handbook of Research on Classroom Assessment* (pp. 125-144). Nueva York, NY: Sage Publications.
- Tierney, R. D., Simon, M. y Charland, J. (2011). Being fair: Teachers' interpretations of principles for standards-based grading. *The Educational Forum*, 75(3), 210-227.
- Torrance, H. (1995). Teacher involvement in new approaches to assessment. En H. Torrance (Ed.), *Evaluating authentic assessment: problems and possibilities in new approaches to assessment* (pp. 44-56). Philadelphia, PA: Open University Press.
- Trigwell, K. (2000). A phenomenographic interview on phenomenography. *Phenomenography*, 1, 62-82.
- Trigwell, K. (1997). Phenomenography: an approach to research. En J. Higgs (Ed.), *Qualitative research: discourse on methodologies* (pp. 39-47). Sidney: Hampton Press.
- Trochim, W. M. (2009). Evaluation policy and evaluation practice. *New Directions for Evaluation*, 123, 13-32.
- Tyler, R. W. (1951). The functions of measurement in improving instruction. En E .F. Lindquist (Ed.), *Educational measurement* (pp. 47-67). Washington, DC: American Council on Education.
- Van den Berg, B. (2002). Teachers' meanings regarding educational practice. *Review of Educational Research*, 72, 577-625.
- Vandeyar, S. y Killen, R. (2007). Educators' conceptions and practice of classroom assessment in post-apartheid South Africa. *South African Journal of Education*, 27(1), 101-115.

- Vey, L. D. (2005). *Enhancing the relationship between learning and assessment*. Tesis doctoral. Canberra, Universidad de Canberra.
- Walker, C. (1998). Learning to learn, phenomenography and children's learning. *Educational and Child Psychology*, 15, 25-33.
- Weiss, R. S. y Rein, M. (1970). The evaluation of broad-aim programs: Experimental design, its difficulties, and an alternative. *Administrative Science Quarterly*, 15(1), 97-109.
- Wiggins, G. (1993). Assessment: Authenticity, context, and validity. *Phi Delta Kappan*, 74, 200-214.
- Winterbottom, A., Bekker, H. L., Conner, M. y Mooney, A. (2008). Does narrative information bias individual's decision making? A systematic review. *Social Science & Medicine*, 67(12), 2079-2088.
- Young, I. M. (2000). *Inclusion and democracy*. Oxford: Oxford University Press
- Young, I. M. (2010). *Responsibility for Justice*. Oxford: Oxford University Press.
- Zabalza, M. Á. (2003). *Competencias docentes del profesorado universitario: calidad y desarrollo profesional*. Madrid: Narcea.



ANEXO 1.

Publicaciones que conforman la tesis

Siguiendo la normativa de la Universidad Autónoma de Madrid sobre las tesis doctorales por compendio de publicaciones, en este primer anexo se presenta el texto completo de los cinco artículos que componen la tesis que aquí se presenta.

A1.1. Publicación 1.
Incidencia de la forma de evaluar los docentes de
Educación Primaria en el rendimiento de los
estudiantes en España

ESTUDIOS
SOBRE
EDUCACIÓN

REVISTA SEMESTRAL DE LA FACULTAD DE EDUCACIÓN Y PSICOLOGÍA



Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España

The Impact of the Evaluation Methodology of the Primary Education Teachers on the Performance of the Student in Spain

F. JAVIER MURILLO

Universidad Autónoma de Madrid
javier.murillo@uam.es

CYNTHIA MARTÍNEZ-
GARRIDO

Universidad Autónoma de Madrid
cynthia.martinez@uam.es

NINA HIDALGO FARRAN

Universidad Autónoma de Madrid
nina.hidalgo@uam.es

Resumen: La forma en cómo se evalúa a los estudiantes en el aula es un elemento fundamental que incide directamente en su desarrollo cognitivo y socio-afectivo. Esta investigación busca determinar la relación entre las actividades de evaluación que llevan a cabo los docentes de Educación primaria de escuelas españolas y el desarrollo en los alumnos de las siguientes competencias: Lingüística, Matemática, Científica, Social y Ciudadana. Para alcanzarlo se realiza una explotación especial de los datos de la Evaluación de Diagnóstico General (EDG) de 2009 a 28.708 estudiantes, 1.341 docentes, 887 centros españoles mediante el uso de Modelos Multinivel de tres de niveles. Los resultados muestran que los estudiantes cuyos docentes realizan en mayor medida controles después de cada tema o lección, corrigen los deberes y cuadernos y valoran el interés y la participación en clase de los estudiantes obtienen mejores resultados. Y, por el contrario, se ha encontrado una relación inversa entre realizar controles escritos trimestrales y el desarrollo de las competencias de los estudiantes.

Palabras clave: evaluación; actividades de evaluación; desarrollo de competencias; enseñanza eficaz; educación primaria; España.

Abstract: The way teachers assess in classroom impact over cognitive and socio-affective development of the students. This paper aims to determine the relationship between the teacher's evaluation to their students, and the academic achievement of their students in four specific areas: Linguistic Competence, Mathematical Competence, Scientific Competence, and Social and Citizenship Competence. To reach our aim, this paper made a special analysis with the data of the Spanish General Diagnostic Evaluation –2009– to 28,708 students, and 1341 teachers of 887 schools. The use of Multilevel Model Analysis with three levels shows the impact of each evaluation strategy on the four competencies studied. Our study confirms those teachers who prepare exams after each lesson, give homework, review notebooks, and value the interest and participation of students in classroom do better. On the other side, we have found an inverse relationship between prepare quarterly exams and the skill development of the students.

Keywords: teaching effectiveness; assessment; primary education; Spain.

INTRODUCCIÓN

Afirmar que las estrategias de evaluación que el docente utiliza en el aula tienen una incidencia en el aprendizaje de sus estudiantes no parece aportar muchas ideas novedosas. Efectivamente, la investigación que busca determinar qué hacen los docentes que consiguen que sus alumnos aprendan más de lo que sería previsible, la investigación sobre enseñanza eficaz ha sido clara en determinar que la forma en que el docente evalúa a sus alumnos, así como su frecuencia y la manera de devolver la información son factores clave asociados al desempeño escolar (Hunt, Wiseman y Touzel, 2009; Muijs y Reynolds, 2000; Orlich, Harder, Callahan, Trevisany y Brown, 2010; Walberg y Paik, 2000).

A pesar de ello, y frente a voces que piden sistemas de evaluación alternativos (p.e. Baker, 2010; Huerta-Macías, 1995; Mateo y Martínez Olmos, 2005), se generaliza la idea de que el simple hecho de evaluar, con la preocupación casi exclusiva de que se garantice la “calidad” psicométrica de los instrumentos, mejorará automáticamente los resultados de los estudiantes. De esta forma, parece ya no importar el qué, el cómo, el cuándo o el para qué se evalúa, parece solo importar que se haga “objetivamente”.

En ese contexto, es necesario dar un paso atrás y volver a mirar cómo evalúan los y las docentes de Educación Primaria en España, y verificar si esa forma de evaluar tiene incidencia en el desempeño de los estudiantes. De esta manera se estará en condiciones de aportar ideas que hagan que se desarrollen sistemas de evaluación que logren realmente incidir en el desarrollo de los estudiantes.

En esta investigación se realizará un Análisis Multinivel de la base de datos de la Evaluación General Diagnóstica de Educación Primaria del año 2009 para determinar qué prácticas educativas tienen incidencia en el aprendizaje de los estudiantes en cuatro materias diferentes.

REVISIÓN DE LA LITERATURA

La investigación sobre Eficacia Escolar, así como sobre Enseñanza Eficaz, ha determinado que las estrategias de evaluación y seguimiento de los estudiantes es un factor que incide de una forma clara directa y medible en el desempeño académico. Efectivamente, ya en el primer estudio sobre Eficacia Escolar desarrollado por Weber (1971), se encontró que la evaluación de los estudiantes era uno de los factores clave que discriminaba entre las buenas escuelas y las malas. Desde entonces, cientos de trabajos han encontrado análogas conclusiones. Así, por ejemplo,

Edmonds (1979) incluyó la evaluación frecuente como uno de los elementos clave que incidían en el aprendizaje de los estudiantes en su famoso “Modelo de 5-factores” que resumía la primera década de investigación sobre eficacia escolar.

A partir de ese momento, diferentes revisiones de estudios llegaron a análogas conclusiones. De los estudios realizados en la década de los 70 dan cuenta Medley y Crook (1980). Tras una concienzuda revisión de más de 600 estudios, estos autores determinaron que la supervisión del trabajo individual de los estudiantes era uno de los cinco comportamientos propios de los docentes eficaces. Black y Wiliam (1998), por su parte, revisaron los trabajos de la década de los 80, y encontraron que la evaluación de los estudiantes, tanto en Educación Primaria como en Secundaria, aumenta los logros de los estudiantes cuando:

- 1 los resultados de la evaluación se utilizan para adaptar el contenido de la enseñanza a los estudiantes,
- 2 los estudiantes reciben retroalimentación que les lleva a mejorar su trabajo,
- 3 tanto docentes como estudiantes conocen los objetivos del aprendizaje y las criterios de evaluación,
- 4 son los propios estudiantes los que pueden evaluarse (auto-evaluación y evaluación entre pares), y
- 5 los estudiantes toman una postura activa en la evaluación.

Teddlie y Reynolds (2000), y posteriormente Sutton (2009), complementaron las características señaladas, apuntando a las razones de los efectos positivos de la evaluación sobre el desarrollo de los estudiantes:

- Genera efectos positivos para el alumno en su autoestima y motivación, dado que siente que el profesor está interesado en su progreso,
- Ayuda a que docentes y estudiantes se centren en los objetivos fundamentales que están siendo evaluados,
- Permite ser utilizado para analizar el progreso del estudiante.
- Contribuye a la mejora de la intervención, aportando datos para valorar su impacto y calidad.

De esta forma, el "para qué evaluar", la finalidad que persiga es un elemento que afecta al rendimiento de los estudiantes. Así lo comprueba el estudio realizado por Fuchs y Fuchs (1986) en el que, a través de un meta-análisis de 21 estudios previos elaborados con estudiantes de Educación Infantil, Primaria y Secundaria, muestra

que en aquellos casos en los que se diseñaba la evaluación con el fin de obtener, sintetizar e interpretar información para facilitar la toma de decisiones orientadas a ofrecer retroalimentación a los estudiantes se genera una mejora estadísticamente significativa sobre su rendimiento académico.

El trabajo elaborado por Crooks (1988) destaca otro elemento clave: la planificación de la evaluación. Así, a través de la revisión de 14 investigaciones previas realizadas a estudiantes de Educación Primaria, Crooks encuentra que la influencia de la evaluación sobre el logro académico, la motivación y el desarrollo de estrategias de aprendizaje en los estudiantes no siempre es positiva, depende de cuán concienzuda haya sido planificada e implementada la evaluación de aula.

La frecuencia con la que se evalúa el progreso de los estudiantes es también un factor que incide en su aprendizaje. Bangert-Drowns, Kulik y Kulik (1991) muestran, a través de un meta-análisis en el que revisa 35 estudios, la relación directa entre la frecuencia de evaluación y la mejora del desempeño de los estudiantes de Educación Primaria. Concretamente, encuentran los efectos positivos de las evaluaciones continuas, aunque también señalan que el aumento de la frecuencia de evaluación no es proporcional al rendimiento, sino que crece muy levemente.

La investigación sobre la estrategia de evaluación más adecuada, sin embargo, no ha dado resultados tan concluyentes, quizá simplemente porque no existe esa "más adecuada". Las investigaciones de Muijs y Reynolds (2001) y de House y Howe (2003), entre otras, sostienen que para potenciar al máximo el desarrollo de los estudiantes y contribuir a la equidad en el aula es preferible recurrir a la combinación de estrategias de evaluación más que al uso de una determinada estrategia en particular. Esa visión de la equidad en el aula es especialmente interesante, dado que no sólo interesa que aprendan los alumnos, sino que lo hagan todos. Resultado análogos fueron hallados por Weurlander, Söderberg, Scheja, Hult y Wernerson (2012). En esa lógica, Stiggins y DuFour (2009) defienden que la elaboración de un sistema evaluador que se ajuste a todos los niveles de aprendizaje de todos los estudiantes requiere de cuatro condiciones: i) diseñada en torno a objetivos claros de aprendizaje; ii) comprometida con la enseñanza basada en el logro de los estándares de aprendizaje; iii) que represente con fiabilidad los objetivos de aprendizaje, y iv) que sus resultados se entreguen de manera oportuna y comprensible.

Un último elemento abordado por la investigación y que queremos destacar es la incidencia de la evaluación no solo en los resultados cognitivos, también en los resultados socio-afectivos de los estudiantes. El reciente estudio longitudinal de cinco años de duración realizado por Kyriakides, Creemers y Antoniou (2009) con

2.505 estudiantes, de 108 aulas localizadas en 50 escuelas de Educación Primaria, clasifica en cinco tipos los comportamientos de los docentes en función de las variaciones encontradas en los resultados escolares (cognitivos y afectivos) de los estudiantes. Una de las principales conclusiones de este estudio es que la evaluación de aula impacta no sólo sobre el desarrollo académico de los estudiantes, sino también en su desarrollo afectivo y psicomotor.

En España, un estudio del Instituto Nacional de Evaluación Educativa (2007) identifica diez factores de la práctica docente asociados al desempeño de los estudiantes. Entre ellos destaca el uso de diferentes formas de evaluación del alumnado de Educación primaria. Así, los docentes dan mucha importancia a la evaluación de la alumna o alumno por sus compañeros de grupo, a la autoevaluación del alumnado y a las pruebas escritas sobre un tema es un factor asociado al rendimiento de los estudiantes. Concretamente llega a explicar el 3% de la varianza. Otros factores relacionados son la Programación de la planificación y evaluación (explican el 6% de la varianza), la revisión de cuadernos y trabajos realizados en casa o en clase y a la corrección individualizada de las actividades o ejercicios (con el 5% de la varianza); el uso del libro de texto y corrección en grupo (el 4% de la varianza) y la observación de las intervenciones del alumnado (5%). De esta forma, se valida la incidencia de la evaluación en el aprendizaje de los estudiantes.

Como complemento de estos datos, en un reciente trabajo elaborado por Vázquez Cano (2012) en 110 centros de Primaria y de Secundaria se encuentra que el 96,5% del profesorado únicamente realiza exámenes de contenido teórico, y que el 92% no devuelve al alumnado una nota desglosada del grado de adquisición de las competencias en la evaluación final. No obstante, los resultados también muestran que el 92% del profesorado no le otorga el peso final de la asignatura a una sola prueba de evaluación, sino que establece porcentajes a las diferentes actividades o instrumentos de evaluación que utiliza en la materia.

Con todo esto se concluye que aunque son numerosas las investigaciones que han evidenciado el impacto de la evaluación sobre el rendimiento académico de los estudiantes a nivel internacional, no son tantos los trabajos que aportan datos a este respecto dentro del panorama nacional español. Esta investigación lleva a cabo un minucioso estudio que busca responder a la siguiente pregunta: ¿las diferentes actividades de evaluación que desarrollan los docentes de Primaria de escuelas españolas inciden en el desarrollo de las competencias académicas del estudiante? ¿qué actividades son las que influyen y qué capacidades son influidas?

METODOLOGÍA

Con esta investigación se pretende determinar la relación entre la forma de evaluar por parte de los docentes de Educación Primaria en España y el desempeño académico de sus estudiantes en cuatro áreas: Competencia Lingüística, Competencia Matemática, Competencia Científica, y Competencia Social y Ciudadana.

Para alcanzar este objetivo se realizó una explotación especial de los datos de la Evaluación de Diagnóstico General (EDG) de Educación Primaria realizada en 2009 por el actual Instituto Nacional de Evaluación Educativa, del Ministerio de Educación de España (Instituto de Evaluación, 2010). Junto con las pruebas de rendimiento en cuatro competencias básicas, se aplicaron una serie de cuestionarios de contexto que aportaron información que ayuda no sólo a contextualizar los resultados de desempeño obtenidos, también permite profundizar en el conocimiento de las escuelas, los docentes y los estudiantes españoles de Educación Primaria.

Las variables utilizadas en el estudio fueron de tres tipos:

1. *Variables de producto.* Desempeño de los estudiantes en cuatro áreas: Competencia en Comunicación Lingüística, Competencia Matemática, Competencia en el Conocimiento y la interacción con el Mundo Físico, y Competencia Social y Ciudadana. Todas ellas estimadas mediante Teoría de Respuesta al Ítem (TRI) y con una puntuación en una escala de media 500 puntos y desviación típica 100.
2. *Variables explicativas.* Forma de evaluar por parte de los docentes, compuesto por ocho variables, todas ellas son variables de escala tipificadas:
 - a. Realizar una prueba sobre conocimientos previos antes de cada tema.
 - b. Realizar controles después de cada tema o lección.
 - c. Realizar controles escritos trimestrales.
 - d. Corregir los deberes y cuadernos.
 - e. Revisar las actividades que hacen en clase.
 - f. Realizar exámenes orales.
 - g. Valorar la autoevaluación del alumnado.
 - h. Valorar el interés y la participación en clase de estudiantes.
3. *Variables de ajuste.* Características del estudiante: Índice socio-económico y cultural de la familia del estudiante (variable tipificada), Género del estudiante (hombre/mujer, variable *dummy*); y país de origen del estudiante (nativo/extranjero, variable *dummy*).

Este estudio hace una explotación especial de la base de datos completa de Evaluación de Diagnóstico General de 2009 para Educación Primaria. Ello significa que utilizó datos de 1.341 docentes de dicha etapa que enseñan en 887 centros educativos, y sus 28.708 estudiantes de 4º curso. El muestreo utilizado por el INE para su obtención, en línea con otras evaluaciones internacionales, fue estratificado por conglomerados, con una estratificación constante de aproximadamente 50 centros en cada una de las Comunidades Autónomas.

Varios fueron los instrumentos que se aplican en la EDG y fueron utilizados por esta investigación. En primer lugar, cuatro pruebas de desempeño una para cada una de las cuatro competencias evaluadas: Competencia en Comunicación Lingüística (fiabilidad de 0,861), Competencia Matemática (0,800), Competencia en el Conocimiento y la Interacción con el Mundo Físico (0,854) y Competencia Social y Ciudadana (0,856). En segundo lugar, un cuestionario dirigido a estudiantes y otro a familias, a partir del cual se elaboró el índice ISEC. Por último, se utilizó un cuestionario para docentes, del cual se obtuvieron los datos sobre su forma de evaluar.

Como estrategia de análisis de datos se utilizaron Modelos Multinivel de tres niveles (estudiante, aula y escuela), como alternativa metodológica más adecuada para abordar datos jerarquizados (Goldstein, 1995; Ma, Ma y Bradley, 2008; Murillo, 2008). Así, se realizaron cuatro de procesos de modelamiento multinivel con el software MLwiN 2.26, uno para cada una de las variables de producto, con los siguientes pasos:

1. Modelo nulo, sólo con la variable producto y el intercepto.
2. Modelo ajustado, con ISEC, género y país de origen.
3. Modelo ajustado con las variables de evaluación introducidas de forma individual.
4. Modelo final, con todas las variables incorporadas de forma simultánea, análogo al siguiente:

$$y_{ijk} = \beta_{0jk} + \beta_{1jk} ISEC_{ijk} + \beta_{2jk} Género_{ijk} + \beta_{3jk} Origen_{ijk} + \beta_{4k} Eval1_{jk} + \dots \\ \dots + \beta_{10k} Eval8_{jk} + e_{ijk}$$

$$\beta_{0jk} = \beta_0 + v_{0k} + u_{0k}$$

$$\beta_{1jk} = \beta_1 + v_{1k} + u_{1k} \dots \beta_{3jk} = \beta_3 + v_{3k} + u_{3k}$$

$$\beta_{4k} = \beta_4 + u_{4k} \dots \beta_{0k} = \beta_0 + u_{0k}$$

Con:

$$[e_{ijk}] \sim N(0, \Omega_e) : \Omega_e = [\sigma_{e_0}^2]$$

$$[u_{jk}] \sim N(0, \Omega_u) : \Omega_u = [\sigma_{u_0}^2]$$

$$[v_k] \sim N(0, \Omega_v) : \Omega_v = [\sigma_{v_0}^2]$$

Donde, para cada i= estudiante en cada j= aula, y k=escuela:

y_{ijk} es la cada una de las cuatro variables de producto: Competencia Lingüística, Competencia Matemática, Competencia Científica y Competencia Social y Ciudadana.

$ISEC_{ijk}$, nivel socio-económico y cultural de la familia del estudiante.

$Género_{ijk}$, género del estudiante (hombre/mujer).

$Origen_{ijk}$, país de origen del estudiante (nativo/inmigrante).

$Eval1_{ij}, \dots, Eval8_{ij}$, cada una de las ocho variables de evaluación del docente estudiadas.

e_{ijk} es el residual para cada estudiante.

u_{jk} el residual de cada aula.

v_k el residual de cada escuela.

RESULTADOS

Tal y como se ha señalado, el proceso de modelización necesario para estudiar la incidencia de las estrategias de evaluación sobre el desempeño de los estudiantes está conformado por cuatro fases consecutivas. En primer lugar, se estiman los cuatro modelos nulos, uno para cada una de las variables de producto (tabla 1). Los mismos sirven como línea base y posibilitan valorar la calidad del modelo.

Tabla 1. Resultados de los modelos multinivel nulos

	COMPETENCIA LINGÜÍSTICA	COMPETENCIA MATEMÁTICA	COMPETENCIA CIENTÍFICA	COMPETENCIA SOCIAL Y CIUDADANA
	β (ee)	β (ee)	β (ee)	β (ee)
PARTE FIJA				
Intercepto	500,31 (1,43)	500,78 (1,40)	500,31 (1,57)	500,49 (1,45)

PARTE ALEATORIA				
Entre escuelas	1386,35 (87,21)	1314,27 (84,92)	1779,13 (102,01)	1462,34 (89,77)
Entre aulas	36,15 (30,26)	43,12 (32,08)	0,00 (0,00)	8,53 (27,27)
Entre estudiantes	8356,74 (77,02)	8529,30 (78,62)	7958,33 (72,76)	8300,08 (76,50)

Fuente: Elaboración propia.

El segundo paso consiste en incluir en cada uno de los modelos nulos las tres variables de ajuste consideradas: ISEC, género y origen del estudiante. En la tabla 2 se presentan los resultados de su estimación. El objetivo de su introducción es controlar su efecto para que no influya en la relación que estamos analizando, es decir, garantizar la validez interna del estudio. Como muestra la tabla 2, el ISEC y el origen del estudiante influyen sobre el desempeño (el origen en tres de las cuatro variables de producto) por lo que es adecuado controlar su efecto.

Además, de estos modelos ajustados se extraen algunas ideas que, aunque alejadas de los objetivos del estudio, resultan interesantes:

- El Índice socio-económico y cultural (ISEC) incide en el desempeño de los estudiantes en las cuatro áreas estudiadas. Así, por cada desviación típica en ISEC por encima (o por debajo) de la media, estudiante obtendrá 28,5 puntos más (o menos) en Competencia Lingüística; 26,5 en Matemáticas; 26,6 en Competencia Científica, y 25,1 en Competencia Social y Ciudadana. Dado que estas variables están escaladas en una media de 500 y una desviación típica de 100, ello significa que por cada desviación típica que suba el índice ISEC de la familia del estudiante, disminuirá o aumentará más de 0,25 desviaciones típicas el desempeño del estudiante en las cuatro medidas de producto.

- El género no incide en el desempeño en ninguna de las cuatro variables analizadas (su aporte no es significativo con un $\alpha=0,05$). Es decir, no hay diferencias significativas en el desempeño de los estudiantes en ninguna de las cuatro variables estudiadas en función de su género.

- Por el contrario, sí existen diferencias significativas en función del Origen del estudiante, de tal forma que los niños y niñas que han nacido fuera de España obtienen resultados más bajos que los estudiantes nativos españoles. En concreto, los datos apuntan a que los estudiantes inmigrantes obtienen 4,11 puntos menos en el desarrollo de la Competencia Lingüística, 4,83 puntos menos en Competencia Científica y 5 puntos menos en Competencia Matemática. Así mismo, los resultados muestran que el origen de los estudiantes no influye sobre el desarrollo de la

Competencia Social y Ciudadana. Por tanto, resulta indiferente ser nativo o inmigrante para el desarrollo Social y Ciudadano del estudiante.

Tabla 2. Resultados de los modelos multinivel ajustados

	COMPETENCIA LINGÜÍSTICA β (ee)	COMPETENCIA MATEMÁTICA β (ee)	COMPETENCIA CIENTÍFICA β (ee)	COMPETENCIA SOCIAL Y CIUDADANA β (ee)
PARTE FIJA				
Intercepto	501,33 (1,16)	501,82 (1,16)	501,32 (1,38)	500,97 (1,27)
ISEC	28,48 (0,64)	26,46 (0,65)	26,58 (0,63)	25,07 (0,65)
Género (niño/niña)	NS	NS	NS	NS
Origen (nativo/ inmigrante)	-4,11 (2,05)	-5,01 (2,01)	-4,83 (2,03)	NS
PARTE ALEATORIA				
Entre escuelas	778,44 (57,35)	786,51 (58,63)	1303,19 (77,86)	1013,48 (67,49)
Entre aulas	33,52 (28,94)	37,63 (30,06)	0,00 (0,00)	7,13 (27,27)
Entre estudiantes	7857,27 (72,42)	8104,70 (74,70)	7487,14 (68,45)	7899,72 (72,81)

Nota: NS Aportación no significativa con $\alpha=0,05$.

Fuente: Elaboración propia.

De los modelos ajustados también es posible estimar el efecto escolar a través del cálculo del Coeficiente de Correlación Intraclase. El coeficiente de correlación intraclase se calcula a partir de las varianzas a nivel escuela, a nivel de aula y a nivel estudiante obtenidas en el proceso de modelización para cada una de las variables de producto según la siguiente fórmula:

$$\rho = \frac{\sigma_{v_0}^2}{\sigma_{v_0}^2 + \sigma_{\mu_0}^2 + \sigma_{\varepsilon_0}^2}$$

Dónde:

$\sigma_{\nu_0}^2$, varianza entre escuelas.

$\sigma_{\mu_0}^2$, varianza entre aulas.

$\sigma_{\mu_0}^2$, varianza entre estudiantes.

Los resultados indican que el efecto escolar es de 8,98% para Competencia Lingüística; 8,81% para Competencia Matemática; 14,82% para Competencia Científica, y 11,36% para Competencia Social y Ciudadana. Cifras análogas se han encontrado en otros estudios (p.e. Murillo, 2005) lo que confirman la adecuación del proceso seguido.

El siguiente paso (tabla 3) es elaborar los modelos multinivel introduciendo individualmente las ocho variables explicativas en cada uno de los cuatro modelos ajustados que se han mostrado en la tabla 2. Con ellos se determinan cuáles son las actividades de evaluación que impactan de manera significativa en el desarrollo de los estudiantes.

TABLA 3. RESULTADOS DE LOS MODELOS MULTINIVEL CONSTRUIDOS A PARTIR DE LA INTRODUCCIÓN DE CADA VARIABLE EXPLICATIVA EN LOS MODELOS AJUSTADOS

	COMPETENCIA LINGÜÍSTICA β (ee)	COMPETENCIA MATEMÁTICA β (ee)	COMPETENCIA CIENTÍFICA β (ee)	COMPETENCIA SOCIAL Y CIUDADANA β (ee)
Realizar una prueba sobre conocimientos previos antes de cada tema	NS	NS	NS	NS
Realizar controles después de cada tema o lección	NS	NS	4,20 (1,61)	4,67 (1,57)
Realizar controles escritos trimestrales	NS	-1,72 (0,80)	NS	-1,68 (0,82)
Corregir los deberes y cuadernos	3,32 (1,16)	NS	4,58 (1,74)	3,92 (1,72)
Revisar las actividades que hacen en clase	NS	NS	4,87 (2,00)	NS
Realizar exámenes orales	NS	NS	NS	NS

Valorar la autoevaluación del alumnado	NS	NS	NS	NS
Valorar el interés y la participación en clase de estudiantes	NS	NS	NS	4,29 (2,08)

Nota: NS Aportación no significativa con $\alpha=0,05$.

Fuente: Elaboración propia.

Del análisis de la tabla 3 se extraen una serie de interesantes ideas para cada una de las áreas evaluadas:

• *Competencia Lingüística*. De las ocho variables explicativas estudiadas relativas a la evaluación, sólo una parece tener efecto en el desempeño en competencia lingüística –descontando el efecto del ISEC y el origen–, el hecho de que el docente corrija los deberes y cuadernos. Concretamente, los datos indican que por cada punto que aumente (o disminuya) sobre la media la variable Corregir los deberes y cuadernos, el desarrollo de la Competencia Lingüística de los estudiantes aumenta (o disminuye) en 3,3 puntos.

• *Competencia Matemática*. De nuevo, una sola variable explicativa parece realizar una aportación estadísticamente significativa al modelo ajustado para esta variable producto: Realizar controles trimestrales. Y, curiosamente, el índice es negativo, lo que significa que el hecho de que el docente realice controles trimestrales tiene una incidencia negativa en el desarrollo de la Competencia Matemática. Expresado en números, por cada punto que el docente aumente (o disminuya) la frecuencia de la realización de controles trimestrales, el desempeño en la variable producto Competencia Matemática de los estudiantes disminuye (o aumenta) 1,72 puntos.

• *Competencia Científica*. En tres de los ocho modelos multinivel calculados con esta variable de producto, uno para cada variable explicativa introducida individualmente, la variable de evaluación hace una aportación estadísticamente significativa, en concreto, con las variables: Corregir los deberes y cuadernos, Realizar controles después de cada tema, y Revisar las actividades realizadas en clase. Ello implica que estas tres variables inciden en el desempeño en la Competencia Científica. Específicamente, los datos muestran que por cada punto que aumente (o disminuya) la frecuencia con la que el docente revisa las actividades realizadas en clase, el estudiante aumenta (o disminuye) en 4,87 puntos su desarrollo en esta competencia, 4,58 por la frecuencia en corregir los deberes y cuadernos, y 4,2 puntos por la frecuencia en realizar controles después de cada tema.

• *Competencia Social y Ciudadana*. Son cuatro las variables explicativas que, según los datos, influyen de manera significativa en el desarrollo de esta competencia: Realizar controles después de cada tema o lección, Realizar controles escritos trimestrales, Corregir los deberes y cuadernos, y Valorar el interés y la participación en clase de estudiantes. Todos ellos con una relación positiva; excepto realizar controles escritos trimestrales que, al igual que ocurría con el desempeño en Competencia Matemática, tiene una relación negativa. En los tres casos positivos la incidencia es de más 4 puntos, y en el caso del coeficiente negativo de 1,7.

Una vez discriminado las variables que realizan una aportación significativa individualmente a los modelos ajustados, para cada variable de producto se estima un modelo con todas ellas de forma simultánea (el modelo final) con el fin de obtener un mayor estimador de máxima verosimilitud en el modelo. El estudio de la variación de la razón de verosimilitud refleja que prácticamente no varía para ninguna de las variables. Así pues, el último paso es el cálculo de los cuatro modelos multinivel final, con las variables de ajuste y todas las variables explicativas de forma simultánea (tabla 4). Para las variables producto Competencia Lingüística y Competencia Matemática, al tener solo una variable explicativa con aportación significativa en los modelos ajustados, estos coinciden con los modelos finales. Luego para estas dos variables estos modelos finales no aportarían más información. Para las otras dos variables de producto, sin embargo, los modelos finales aportan información sobre las variables que inciden más en el desarrollo de las competencias, y las que hacen una aportación diferente.

Tabla 4. Resultados de los modelos multinivel finales

	COMPETENCIA LINGÜÍSTICA β (ee)	COMPETENCIA MATEMÁTICA β (ee)	COMPETENCIA CIENTÍFICA β (ee)	COMPETENCIA SOCIAL Y CIUDADANA β (ee)
PARTE FIJA				
Intercepto	492,33 (4,73)	503,57 (1,45)	481,13 (1,60)	479,34 (6,96)
ISEC	28,64 (0,69)	26,44 (0,65)	26,63 (0,63)	25,12 (0,68)
Origen (nativo-inmigrante)	-4,10 (2,01)	-5,04 (2,10)	-4,75 (2,03)	NS
VARIABLES DE PROCESO				
Realizar controles después de cada tema o lección			3,73 (1,62)	4,32 (1,16)

Corregir los deberes y cuadernos	3,32 (1,16)		4,26 (2,00)	
Realizar controles escritos trimestrales		-1,67 (0,80)		-1,18 (0,82)
Valorar el interés y la participación en clase de estudiantes				4,34 (2,10)
PARTE ALEATORIA				
Entre escuelas	774,88 (57,18)	785,66 (58,46)	1277,23 (76,56)	987,30 (66,48)
Entre aulas	33,35 (28,93)	35,84 (29,94)	0,00 (00,00)	5,94 (27,16)
Entre estudiantes	7856,91 (72,42)	8104,61 (74,70)	7488,80 (68,46)	7878,70 (72,80)

Nota: NS Aportación no significativa con $\alpha=0,05$.

Fuente: Elaboración propia.

Analizando la tabla 4 se observa que, para Competencia Científica, de las tres variables que tuvieron una aportación significativa individualmente se “cae” la variable Revisar las actividades que realizan en casa. Para la variable producto Competencia Social y Ciudadana de las cuatro variables que realizan aportaciones en los modelos individuales, y tomadas conjuntamente solo tres de ellas realiza una aportación significativa; se pierde, por tanto, la variable corregir los deberes.

Con todo ello, se puede concluir que las actividades de evaluación que realiza el docente en el aula tiene una incidencia en el desarrollo global del estudiante controlando el efecto del nivel socio-económico y el origen (tabla 5). Más exactamente, tres son las actividades evaluativas que han mostrado su relación positiva con el desempeño en alguna o en varias de las competencias estudiadas:

- Realizar controles después de cada tema o lección, que incide sobre el desarrollo de su Competencia Científica y en el desarrollo de su Competencia Social y Ciudadana.
- Corregir los deberes y cuadernos, que impacta sobre el desempeño de los estudiantes en Competencia Lingüística y en Competencia Científica.
- Valorar el interés y la participación de los estudiantes en clase, que afecta de manera positiva sobre el desarrollo de su Competencia Social y Ciudadana.

Por el contrario, llevar a cabo controles escritos trimestrales afecta negativamente en el desarrollo de los estudiantes en Competencia Matemática y en Competencia Social y Ciudadana.

Por último, hay tres actividades de evaluación que según estos datos no parecen estar directamente asociadas con el desarrollo de ninguna competencia:

- Realizar una prueba sobre conocimientos previos antes de cada tema.
- Realizar exámenes orales.
- Valorar la autoevaluación del alumnado.

Tabla 5. Resumen del proceso de modelado multinivel

	COMPETENCIA LINGÜÍSTICA	COMPETENCIA MATEMÁTICA	COMPETENCIA CIENTÍFICA	COMPETENCIA SOCIAL Y CIUDADANA
Realizar una prueba sobre conocimientos previos antes de cada tema				
Realizar controles después de cada tema o lección			XX	XX
Realizar controles escritos trimestrales		XX*		XX*
Corregir los deberes y cuadernos	XX		XX	X
Revisar las actividades que hacen en clase			X	
Realizar exámenes orales				
Valorar la autoevaluación del alumnado				
Valorar el interés y la participación en clase de estudiantes				XX

Nota: X variables con una aportación significativa en los modelos con variables introducidas individualmente. XX variables con una aportación significativa en los modelos finales

(*) aportación negativa

Fuente: Elaboración propia.

Vista la incidencia de algunas actividades de evaluación sobre el desarrollo de las diferentes competencias, puede ser interesante analizar qué es lo que hacen los docentes de Educación Primaria en España en las otras variables sobre diferentes acti-

vidades de evaluación. De esta forma será posible valorar en qué medida las prácticas evaluativas que se han mostrado especialmente eficaces son desarrolladas por los maestros y las maestras en España. Los datos globales aparecen en la tabla 6.

Tabla 6. Distribución de los docentes según la frecuencia de realización de diferentes actividades de evaluación. Datos en porcentaje

	NUNCA O CASI NUNCA	ALGUNAS VECES	CASI SIEMPRE	SIEMPRE
Realizar una prueba sobre conocimientos previos antes de cada tema	31,15	37,78	22,01	9,06
Realizar controles después de cada tema o lección*	,90	4,51	23,67	70,92
Realizar controles escritos trimestrales*	45,57	22,52	11,96	19,95
Corregir los deberes y cuadernos*	,08	2,18	23,98	73,76
Revisar las actividades que hacen en clase	,08	1,13	24,00	74,79
Realizar exámenes orales	12,80	62,58	17,92	6,70
Valorar la autoevaluación del alumnado	12,76	42,14	28,40	16,70
Valorar el interés y la participación en clase de alumnas/os*	,00	1,20	16,69	82,11

Notas: en negrita las modas.

(*) variables asociadas con el desempeño

Fuente: Elaboración propia a partir de los datos de le EGD 2009.

Los datos aportados por los 1.341 docentes estudiados muestran con claridad las actividades que desarrollan respecto a la evaluación de sus estudiantes. En una rápida mirada es posible agrupar las prácticas en tres grandes grupos en función de su frecuencia de uso.

Por una parte están las prácticas de evaluación desarrolladas por la mayoría de los docentes de Primaria, que serían cuatro: Realizar controles después de cada tema o lección, Corregir los deberes y cuadernos, Revisar las actividades hechas en clase, y Valorar para la calificación el interés y la participación en clase de los estudiantes. En todas ellas, más del 70% de los maestros y maestras afirman hacerlo siempre (el 70,9% Realizar controles después de cada tema o lección, el 73,7% Corregir los deberes y cuadernos, y el 74,8% valorar para la calificación el interés y la partici-

pación en clase de los estudiantes), y más del 96% de ellos dicen hacerlo siempre o casi siempre (Realizar controles después de cada tema o lección el 96,6%, Corregir los deberes y cuadernos el 97,7%, Revisar las actividades hechas en clase el 98,8%, y Valorar el interés y la participación en clase de los estudiantes el 98,8%).

En el segundo grupo estarían las actividades de evaluación realizadas de forma poco habitual. Allí se encuentran tres de las actividades analizadas: Realizar pruebas sobre conocimientos previos antes de cada tema, Realizar pruebas orales, y Tener en consideración la propia autoevaluación del alumnado. En los tres casos, la puntuación modal estaría en la categoría “algunas veces” y en todos ellos hay más docentes que afirman que no las han desarrollado “nunca” o “algunas veces” que aquellos que dicen desarrollarlas “casi siempre” o “siempre”. Concretamente, el porcentaje de docentes que afirma Desarrollar pruebas sobre conocimientos previos antes de cada tema es del 63,9%, realizar exámenes orales del 75,3%, y Valorar la autoevaluación del alumnado un 55,0%.

Por último se encuentra la actividad Realizar controles escritos trimestrales, en la que el 45,6% de los maestros y maestras dicen no realizarla nunca. Esto es interesante porque es la actividad que se había encontrado que tenía un efecto negativo sobre el aprendizaje de los estudiantes.

DISCUSIÓN Y CONCLUSIONES

Los resultados obtenidos en esta investigación muestran, en primer lugar, que la forma de evaluar por parte de los y las docentes de Educación Primaria en España está directamente relacionada con el desempeño académico de los estudiantes en las cuatro áreas analizadas: Competencia Lingüística, Competencia Matemática, Competencia Científica, y Competencia Social y Ciudadana. Estos resultados, por sí solos, tienen la valía de confirmar –con datos actuales, en nuestra realidad y con una muestra estadísticamente representativa para todo el país– un hecho ya demostrado por la experiencia y la investigación (p.e. Black y Wiliam, 1998; Stiggins y DuFour, 2009; Sutton, 2009; Teddlie, 2000): el cómo son evaluados incide en cuánto aprenden los estudiantes. Este hallazgo, sin embargo, no por conocido es menos importante; subraya la necesidad de mejorar los sistemas de evaluación que el docente usa para que los estudiantes aprendan más y mejor.

El análisis detallado acerca de qué actividades de evaluación están asociadas al desempeño aporta resultados más interesantes, novedosos y útiles. De las ocho actividades evaluativas estudiadas por la Evaluación de Diagnóstico General (EDG), que desarrollan los docentes para Primaria, cinco han mostrado su relación con el

desempeño en alguna competencia.

En primer lugar, se ha observado que realizar controles después de cada tema o lección incide positivamente en el aprendizaje de los estudiantes. Este resultado hace referencia a la importancia de tener la enseñanza bien estructurada en lecciones, y en las que en cada una estén perfectamente delimitados los objetivos, los contenidos, las actividades y las actividades de evaluación de carácter formativo (Bangert-Drowns, Kulik y Kulik, 1991; Hunt, Wiseman y Touzel, 2009; Killen, 2005; Murillo, Martínez-Garrido y Hernández-Castilla, 2011).

Un docente que realiza evaluaciones al final de cada tema es un docente que hace evaluaciones frecuentes. Y la investigación ha encontrado relación entre la evaluación frecuente y el desarrollo cognitivo del estudiante (Muijs, Chapman, Collins y Armstrog, 2010; Murillo, 2007a, 2007b), pero también con el no desarrollo socio-afectivo (Kyriakides, Creemers y Antoniou, 2009; Murillo y Hernández-Castilla, 2011). En esa lógica puede también entenderse otro de los hallazgos de este estudio: los estudiantes de los docentes que realizan controles trimestrales escritos obtienen peores resultados en comparación con los que no los realizan. Seguramente porque los maestros y maestras que realizan controles al final del trimestre no realizan estas evaluaciones frecuentes de las que hablamos.

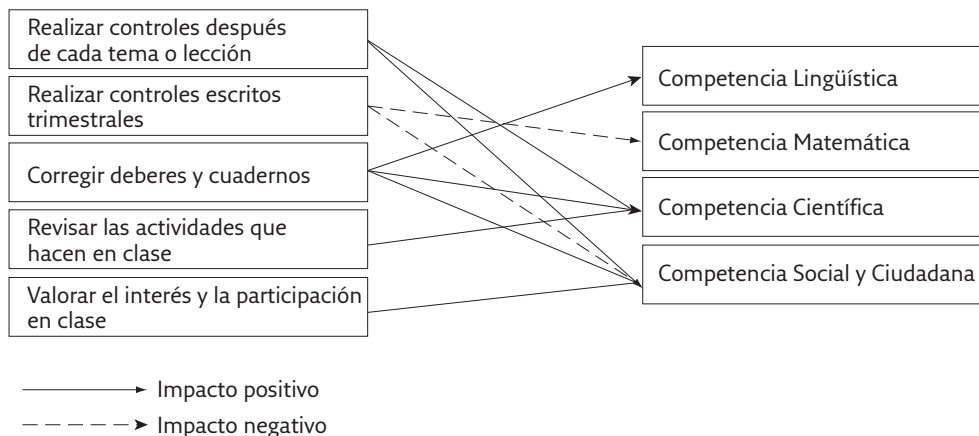
Una actividad evaluativa más que, según esta investigación, incide positivamente en el desempeño de los estudiantes, es que el docente corrija los deberes y cuadernos; al igual que revisar las actividades que hacen en clase. Estos resultados son coherentes con los encontrados por Carlson y Francis (2002). Estos investigadores encontraron que cuando los docentes corrigen las actividades realizadas por los estudiantes y se las devuelven hay un aprendizaje significativamente mayor de éstos.

Especialmente interesante es el hallazgo que establece una relación directa y medible entre la valoración del interés y la participación en clase de los estudiantes y su aprendizaje. Otro estudio anterior, de carácter experimental, realizado por Vey (2005), refleja la mejora del rendimiento de los estudiantes cuando la evaluación considera la colaboración y comunicación entre los propios estudiantes.

Esta investigación, sin embargo, no ha encontrado relación entre Realizar una prueba sobre conocimientos previos antes de cada tema, Realizar exámenes orales y Valorar la autoevaluación del alumnado, y el desarrollo de competencias por parte de los estudiantes. Estos resultados no son coincidentes con otras investigaciones desarrolladas en el ámbito internacional (ver, por ejemplo, Muijs et al., 2010). Habrá que seguir investigando las razones de esas discrepancias.

En el gráfico 1 se ha representado las relaciones encontradas entre actividades de evaluación desarrolladas y las competencias en las que impactan, una de las preguntas de investigación formuladas.

Gráfico 1. Interrelación entre la actividad de evaluación que desarrolla el docente y la competencia en la que impacta



Fuente. Elaboración propia.

El estudio descriptivo de los que hacen (o dicen hacer) los maestros y maestras de Educación Primaria respecto a esos elementos estudiados permite una valoración de hasta qué punto se están haciendo bien o mal las cosas. Y los datos indican que las cosas se hacen mayoritariamente bien. Una importante mayoría de los docentes realiza controles después de cada tema, corrige deberes y cuadernos, valora el interés y la participación en clase de los estudiantes y revisa las actividades en clase. Por el contrario, son proporcionalmente pocos los que hacen controles escritos trimestrales, elemento que parece tener una incidencia negativa en el aprendizaje.

En todo caso, la principal lección aprendida de esta investigación es la necesidad de cuidar la evaluación como un elemento fundamental del proceso de enseñanza y aprendizaje. Así, si queremos mejorar el desarrollo integral de los estudiantes, no sólo es importante evaluar, sino cómo es esa evaluación. Una evaluación inadecuada, como hemos visto, puede incluso entorpecer el desarrollo global de los estudiantes. De ahí que sea necesario ponerlo en primera línea de atención en la formación tanto inicial como permanente de los docentes.

La principal fortaleza metodológica de esta investigación es el tamaño de la muestra (1.341 docentes de Educación Primaria) y la estrategia de muestreo utilizada (estratificado por conglomerados), lo que permite garantizar la validez externa de

los resultados; así como las técnicas de análisis de datos, los Modelos Multinivel, sin duda los más adecuados para contrastar datos de diferentes niveles (escuela, aula y estudiante), lo que contribuye a garantizar la validez interna, así como a obtener las estimaciones más ajustadas. Las limitaciones vienen de la fuente de datos: los cuestionarios para docentes. En dos sentidos, por un lado, al no haber sido contruados específicamente para estudiar las actividades de evaluación de los docentes, hay muchas actividades de las que no tenemos información, por ejemplo el uso de diferentes estrategias metodológicas o la frecuencia de evaluación. Y, por otro, la limitación propia de saber sólo lo que dice el docente, no lo que ocurre realmente en el aula.

Todo ello lleva inexorablemente a plantearse la necesidad de más investigaciones sobre el tema. En esa idea sería importante que el Instituto de Evaluación siga mantenido esas preguntas en futuros estudios, pero también que lo complete con más variables relevantes, por ejemplo las ya comentadas sobre la frecuencia de la evaluación y la retroalimentación, y sobre las estrategias de evaluación seguidas y sobre su utilización. En todo caso, parece importante realizar un estudio observacional para ver la incidencia de las estrategias de evaluación en el desempeño de los estudiantes.

Stiggins y DuFour (2009) defendieron la evaluación como una de las herramientas educativas más poderosas que tanto maestros como escuelas tienen a su disposición para alcanzar el logro de los estudiantes. Los resultados de este estudio así lo confirman, de tal forma que podemos hablar de la evaluación como el “combustible que permite iniciar el camino hacia la transformación y mejora de las escuelas. Un camino cada día más necesario.

Fecha de recepción del original: 14 de mayo de 2013

Fecha de aceptación de la versión definitiva: 14 de febrero de 2014

REFERENCIAS

- Baker, E. (2010). *What Probably Works in Alternative Assessment*. Los Ángeles: CRESST, UCLA.
- Bangert-Drowns, R. L., Kulik, J. A. y Kulik, C. C. (1991). Effects of frequent classroom testing. *Journal of Educational Research*, 85, 89-99.
- Black, P. y Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7-24.
- Carlson, C. D. y Francis, D. J. (2002). Increasing the reading achievement of at-

- risk children through direct instruction: evaluation of the Rodeo Institute for Teacher Excellence. *Journal of Education for students placed at risk*, 7(2), 141-166.
- Crooks, T. J. (1988). The Impact of Classroom Evaluation Practices on Students. *Review of Educational Research*, 58(4), 438-481.
- Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37(1), 15-24.
- Fuchs, L. S. y Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, 53(3), 199-208.
- Goldstein, H. (1995). *Multilevel Statistical Models*. Nueva York: Willey.
- House, E. y Howe, K. (2003). Democratic deliberative evaluation. En T. Kellaghan, D. Stufflebeam y L. Wingate (Eds.), *The International Handbook of Evaluation Studies* (pp. 79-100). Dordrecht: Kluwer.
- Huerta-Macías, A. (1995). Alternative Assessment: Responses to Commonly Asked Questions. *TESOL Journal*, 5(1), 8-11.
- Hunt, G. H., Wiseman, D. G. y Touzel, T. J. (2009). *Effective teaching: preparation and implementation*. Springfield: Charles C. Thomas.
- INEE-Instituto Nacional de Evaluación Educativa (2007). *Evaluación del Sistema Educativo en España en Educación Primaria*. Madrid: INEE.
- Instituto de Evaluación (2010). *Evaluación General de Diagnóstico 2009*. Madrid: Instituto de Evaluación.
- Killen, R. (2005). *Programming and assessment for quality teaching and learning*. Sidney: Thomsom.
- Kyriakides, L., Creemers, B. P. M. y Antoniou, P. (2009). Teacher Behaviour and Student Outcomes: Suggestions for Research on Teacher Training and Professional Development. *Teaching and Teacher Education*, 25, 12-23.
- Ma, X., Ma, L., y Bradley, K. D. (2008). Using multilevel modeling to investigate school effects. En A. A. O'connell y B. McCoach (Eds.), *Multilevel Modeling of Educational Data* (pp. 59-110). Nueva York: Information Age Publishing Inc.
- Mateo, J. y Martínez Olmo, F. (2005). *L'avaluació alternativa dels aprenentatges*. Barcelona: ICE Universitat de Barcelona.
- McMillan, J. (1997). *Classroom Assessment. Principles and Practices for Effective Instruction*. Needham Heights: Allyn & Bacon.
- Medley, D. y Crook, P. R. (1980). Research in Teaching Competency and Teaching Tasks. *Theory into practice*, 19(4), 293-316.
- Muijs, D. y Reynolds, D. (2000). School Effectiveness and Teacher Effectiveness in Mathematics. *School Effectiveness and School Improvement*, 11(3), 273-303.
- Muijs D. y Reynolds D. (2001). *Effective teaching. Evidence and practice*. Londres: Paul Chapman Publication.

- Muijs, D., Chapman, C., Collins, A. y Armstrog, P. (2010). *Maximum Impact Evaluation*. Manchester: University of Manchester.
- Murillo, F. J. (2005). ¿Importa la escuela? Una estimación de los efectos escolares en España. *Tendencias Pedagógicas*, 10, 29-45.
- Murillo, F. J. (2007a). School Effectiveness Research in Latin America. En T. Townsend (Ed.), *International Handbook of School Effectiveness and Improvement* (pp. 75-92). Nueva York: Springer.
- Murillo, F. J. (2007b). Investigación iberoamericana sobre *eficacia escolar*. Bogotá: Convenio Andrés Bello.
- Murillo, F. J. (2008). Los Modelos Multinivel como herramienta para la investigación educativa. *Magis. Revista Internacional de Investigación Educativa*, 1(1), 17-34.
- Murillo, F. J. y Hernández-Castilla, R. (2011). Factores escolares asociados al desarrollo socio-afectivo en Iberoamérica. *RELIEVE*, 17(2), art 2.
- Murillo, F. J., Martínez-Garrido, C. y Hernández-Castilla, R. (2011). Decálogo para una enseñanza eficaz. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 9(1), 6-27.
- Orlich, D. C., Harder, R. J., Callahan, R. C., Trevisan, M. S. y Brown, A. H. (2010). *Teaching Strategies: A Guide to Effective Instruction*. Boston: Wadsworth.
- Reynolds, D. y Teddlie, C. (2000). *The international handbook of school effectiveness research*. Londres: Falmer Press.
- Stiggins, R. y DuFour, R. (2009). Maximizing the power of formative assessment. *Phi Delta Kappan*, 90(9), 640-644.
- Sutton, R. (2009). Formative Assessment the Way the School Does Business: The impact and Implications of formative assessment for teachers, students and school leaders. En A. Hargreaves, A. Lieberman, M. Fullan y D. Hopkins (eds.), *Second International Handbook of Educational Change* (pp. 883-899). Londres: Springer.
- Teddlie, C. y Reynolds, D. (2000). *The International Handbook on School Effectiveness Research*. Nueva York: Routledge Falmer.
- Vázquez Cano, E. (2012). La evaluación del aprendizaje en primaria y secundaria: los indicadores de evaluación. *Espiral: Cuadernos del Profesorado*, 5(10), 30-41.
- Vey, L. D. (2005). *Enhancing the relationship between learning and assessment*. Tesis doctoral. Canberra: University of Canberra. Extraído el 26 de junio de 2014 de http://www.canberra.edu.au/researchrepository/file/edc0ca80-acb7-25b9-4fa8-6e03a01d4e27/1/full_text.pdf
- Weber, G. (1971). *Inner-city children can be taught to read: four successful schools*. Washington: Council for Basic Education.

- Weurlander, M., Söderberg, M., Scheja, M., Hult, H. y Wernerson, A. (2012). Exploring formative assessment as a tool for learning: students experiences of different methods of formative assessment. *Assessment and Evaluation in Higher Education*, 37(6), 747-460.
- Walberg, J. H. y Paik, D. J. (2000). *Effective educational practices*. Ginebra: IBE/ UNESCO.

A1.2. Publicación 2. **Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes**

RINACE

Revista Iberoamericana sobre
Calidad, Eficacia y Cambio en Educación

ISSN: 1696-4713

REICE

**Revista Iberoamericana sobre Calidad,
Eficacia y Cambio en Educación**

Enero 2017 – Volumen 15, número 1
doi: 10.15366/reice2017.15.1

rinace.net/reice/
revistas.uam.es/reice



Las Concepciones sobre el Proceso de Evaluación del Aprendizaje de los Estudiantes

Conceptions about Assessment Process of Students' Learning

Nina Hidalgo *
F. Javier Murillo

Universidad Autónoma de Madrid

La evaluación de los estudiantes es una de las prácticas docentes que más inciden en el aprendizaje y desarrollo de los estudiantes, especialmente porque determina su vida presente y futura. Numerosos estudios han evidenciado la relación entre cómo evalúan los docentes y sus concepciones sobre evaluación. Fruto de su importancia, el estudio de las concepciones de evaluación va adquiriendo fuerza en las últimas décadas. Por ello, el presente estudio busca profundizar en los estudios actuales que se centran en conocer las concepciones sobre evaluación. Para lograr dicho objetivo hemos estructurado el artículo en tres partes. En primer lugar, ahondamos en la definición de las concepciones y su concreción en el ámbito educativo y evaluativo. En segundo lugar, revisamos las investigaciones de concepciones sobre evaluación en cuatro direcciones: a) los tipos de concepciones, b) los factores que inciden en su diversidad, c) la relación entre las concepciones y la práctica evaluativa y d) otras investigaciones sobre concepciones. Por último, cerramos el presente trabajo con unas reflexiones acerca de lo que hemos aprendido de los estudios analizados relacionados con las concepciones sobre evaluación así como la propuesta de futuras líneas de investigación centradas en aquello que aún nos queda por saber de las creencias implícitas sobre la evaluación del aprendizaje de los estudiantes.

Descriptores: Evaluación de Estudiantes, Concepciones, Concepciones sobre evaluación, Práctica evaluativa

Students' Assessment is one of the teaching practices which most affect into learning and development of students, especially because it determines their present and future life. Numerous studies have demonstrated the relationship between how teachers assess and their conceptions about teacher' assessment. As a result of its importance, the study of assessment' conceptions has gained strength in recent decades. Therefore, the present study seeks to deepen into the present studies that focus on knowing the conceptions about assessment. To achieve this goal, we have structured the paper into three parts. Firstly, we delve into the definition of conceptions and their concreteness in the educational and evaluative field. Secondly, we review the group of researches on assessment' conceptions in four directions: a) types of conceptions, b) factors that affect this diversity, c) the relationship between conceptions and assessment practice, and d) other studies related to conceptions. Finally, we close the present work with some personal reflections about what we have learned from the studies analyzed related to assessment' conceptions as well as to propose future lines of research centered on what still remains to be known about implicit beliefs of students' learning' assessment.

Keywords: Students' assessment, Conceptions, Assessment' conceptions, Assessment practice.

*Contacto: nina.hidalgo@uam.es

1. Introducción

No es muy arriesgado afirmar que la evaluación de los estudiantes es una de las actividades con peor reputación del sistema educativo. Al menos así lo aseveran investigaciones como la realizada por Sambel, McDowell y Brown (1997). Estos autores encontraron que docentes y estudiantes perciben de forma muy negativa la práctica de la evaluación, tanto por cómo se desarrolla el proceso evaluativo como por sus implicaciones y la relacionaron con una evaluación entendida esencialmente como rendición de cuentas, donde la práctica evaluativa sirve para certificar, clasificar y etiquetar a los estudiantes.

Quizá por ello, o quizá por la importancia de la evaluación en el proceso de enseñanza y aprendizaje y la influencia de las concepciones sobre su desarrollo, en los últimos años ha surgido con fuerza una línea de investigación empírica que busca profundizar en el estudio de las concepciones de evaluación tanto de docentes como de estudiantes. El motivo es claro: en la medida que conozcamos y comprendamos mejor las concepciones que tienen los estudiantes sobre la evaluación, estaremos en condiciones de desarrollar mejores evaluaciones, aquellas que se ajusten más a sus necesidades y expectativas, y con ello, mejorar su aprendizaje (Brown, 2008; Murillo e Hidalgo, 2015).

Una de las líneas de investigación que está tomando más fuerza en la actualidad es aquella que busca examinar las concepciones sobre evaluación de estudiantes que se desarrolla en el aula (p. ej., Coll y Remesal, 2009; Griffiths, Gore y Ladwig, 2006; Harris y Brown, 2009; Pope, Green, Johnson y Mitchell, 2009; Remesal, 2011). La idea que sustenta el auge de esta línea de investigación es clara: dado que las concepciones son las ideas previas que preceden a la acción, sólo conociendo y cambiando las concepciones podremos mejorar la práctica evaluativa y, con ello, el aprendizaje y desarrollo integral de los estudiantes (Griffiths, Gore y Ladwig, 2006; Murillo, Martínez-Garrido e Hidalgo, 2014).

En el presente artículo se hace una revisión de los avances en la investigación acerca de las concepciones de docentes y estudiantes sobre la evaluación del aprendizaje. Para ello se ha organizado en tres grandes bloques. En primer lugar se define qué son las concepciones y cómo se aplican a la educación y a la evaluación educativa, a continuación se describe el propósito, la metodología y los resultados de los principales estudios centrados en las concepciones sobre evaluación y, por último, se presenta una reflexión final proponiendo estudios para el futuro sobre esta línea de investigación.

2. Una mirada hacia las concepciones sobre evaluación

Desde la perspectiva de la psicología cognitiva, las concepciones se entienden como un elemento esencial para comprender la psique y el comportamiento humano, ya que hacen referencia a las representaciones de las personas del mundo que les rodea (Pozo, Scheuer, Mateos y Pérez-Echeverría, 2012). En este sentido, las concepciones se definen como un sistema organizado de creencias, originadas en las experiencias del individuo y desarrolladas a través de las interacciones en las que participa (Coll y Remesal, 2009; Hasselgren y Beach, 1997; Martín, Mateos, Martínez, Cervi, Pecharromán y Villalón, 2006; Marton, 1981; Remesal, 2011; Säljö, 1994).

Las personas, por naturaleza, construimos nuestro pensamiento a partir de ideas, opiniones y creencias originada en nuestras propias experiencias. En función de las vivencias, cada ser humano concibe y construye en su mente el mundo de forma distinta al resto de personas, y esto incide en la forma en que se comporta e interacciona a lo largo de su vida. Por lo tanto, aquellas concepciones que tenga cada persona determina en gran medida la forma en que actúe y la intencionalidad en dichos actos (Atkinson y Claxton, 2000; Harris, 2008; Murillo, Hidalgo y Flores, 2016; Pozo, 2006).

A lo largo de la historia, se han utilizado términos similares como teorías implícitas, suposiciones o creencias razonables para referirse a las concepciones. No obstante, investigadores como Thompson (1992) defienden que es necesario un significado para concepciones diferente y específico entendiéndolo como una “estructura mental más general que abarca creencias, significados, conceptos, proposiciones, reglas, imágenes mentales, preferencias, etc. que configuran la forma en que las personas nos enfrentamos a distintos fenómenos” (p. 130). De acuerdo con esto último, Barnes, Fives y Dacey (2015) y Philipp (2007) confirman la importancia de utilizar el término de concepciones, más amplio y adecuado para todas las creencias humanas que afectan a la visión del mundo.

De acuerdo con Pozo, Scheuer, Mateos y Pérez-Echeverría (2006) las representaciones implícitas o concepciones de las personas tienen supremacía sobre sus actos, siendo la principal herramienta de aprendizaje y acción. Por ello, para poder cambiar las prácticas de las personas es fundamental conocer previamente sus concepciones y vincularlas a su realidad.

Una de las características principales de las concepciones es su carácter eminentemente social. Así, lejos de ser individuales, las concepciones nacen y se construyen en entornos sociales, ya que se fundamentan e interrelacionan en espacios y situaciones culturalmente compartidos. En este sentido, nuestras concepciones se construyen en la interacción con otras personas, por lo que la forma en que nosotros concibamos el mundo está influido directamente por los individuos que nos rodean (Pozo, 2006; Van den Berg, 2002).

Centrándonos en el ámbito educativo, las concepciones de los profesores y profesoras son una red de creencias, ideas y opiniones que influyen directamente en la forma en que éstos entienden el proceso de enseñanza-aprendizaje e interactúa diariamente con sus estudiantes y el resto de compañeros docentes (Griffiths, Gore y Ladwig, 2006; Prieto y Contreras, 2008; Thompson, 1992). Las concepciones docentes se entienden, por tanto, como un marco o estructura de ideas implícitas personales a través de las cuales los docentes interpretan su práctica profesional (Philipp, 2007; Pratt, 1992; Thompson, 1992). De esta forma, las concepciones del profesorado son influidas por el contexto en el que se desarrolla el profesor, definiendo su práctica docente. Queda claro que dicha influencia no es siempre coherente o explícita, pero sin duda las concepciones son un factor decisivo que determina el proceso de enseñanza-aprendizaje (Ekeblad y Bond, 1994; Griffiths, Gore y Ladwig, 2006).

No obstante, aunque las concepciones definan la práctica profesional de los docentes, estas creencias no son conformadas de forma explícita y consciente por parte del profesorado. Las experiencias de vida, el contexto en el que trabajan, sus experiencias previas como estudiantes del sistema educativo y las presiones sociales y políticas configuran estas ideas implícitas de los docentes, integrándolas en sus estructuras

cognitivas e influyendo directamente en su forma de comprender la enseñanza, el aprendizaje o la evaluación (Prieto y Contreras, 2008).

En la psique de los profesores y profesoras no todas las concepciones tienen el mismo peso. Así, existen creencias centrales y otras complementarias que ayudan a matizar las más importantes (Pajares, 1992; Thompson, 1992). Dentro de estas concepciones, tanto generales como periféricas, existen algunas más conceptuales y otras más centradas en el desempeño diario docente (Marton, 1981), convirtiéndose en una estructura en forma de red donde las concepciones están interconectadas.

Si trasladamos el concepto de concepciones a la evaluación, se entienden como aquellas ideas previas, creencias, teorías implícitas y estructuras mentales que dibujan la forma en la que los profesores y profesoras entienden la evaluación de los estudiantes así como su ejecución (Brown, 2004; Scott, Webber, Lupart, Aitken y Scott, 2014). La evaluación del alumnado puede ser concebida de distintas formas, existiendo múltiples concepciones de evaluación, las cuales, al tener un carácter social, suelen ser influidas por muchos factores, incluyendo el contexto y la cultura en la que trabajan, su formación previa o sus experiencias escolares vividas (Brown, 2003; Rubie-Davies, 2015; Scott et al., 2014). En este sentido, la forma en que evalúa un docente refleja, en parte, sus ideas implícitas construidas en interacción con otros sujetos sobre cómo entiende la evaluación, impactando directamente en su práctica evaluativa (Marshall y Drummond, 2006).

3. Las investigaciones sobre concepciones de evaluación. Una revisión actual

Tal y como venimos diciendo, un fructífera y reciente línea de investigación es la que estudia las concepciones sobre el proceso de evaluación del aprendizaje (Atjonen, 2014; Azis, 2012; Brown, 2003, 2004; Brown y Hirschfeld, 2008; Brown y Remesal, 2012; Brown, Lake y Matters, 2011; Chen, Morris y Mansour, 2012; Fives y Gill, 2014; Harris y Brown, 2008, 2009; Winterbottom et al., 2008; entre otros).

Uno de los primeros estudios centrado en las concepciones sobre evaluación fue el del profesor estadounidense M. Frank Pajares (1992). En su estudio, Pajares (1992) confirmó que profundizar en las concepciones de la enseñanza es prioritario para la investigación en educación, ya que sólo si conocemos las concepciones podemos cambiar las prácticas del profesorado en el aula. En este sentido, conociendo las creencias implícitas podremos determinar las prácticas docentes y así trazar líneas para su mejora, fin último de cualquier investigación educativa (Brown, 2008; Opre, 2010; Pajares, 1992; Sambell, McDowell y Brown, 1997; Vandeyar y Killen, 2007).

El interés evidente mostrado por la investigación en las concepciones sobre evaluación parte de la premisa de que estas concepciones, al igual que otros tipos de creencias de los profesores, influyen significativamente en sus decisiones y actividad profesional (Brown, 2008; Opre, 2010; Vandeyar y Killen, 2007). Por lo tanto, si queremos comprender cómo evalúan los profesores es fundamental conocer sus concepciones. Con esta lógica de fondo, nos interesa conocer las distintas direcciones que han cogido las investigaciones centradas en las concepciones sobre evaluación para aprender de ellas.

Para acercarnos a estas diferentes investigaciones sobre concepciones de evaluación que existen, hemos considerado fundamental centrarnos en aquellos estudios a) focalizados

en identificar las distintas concepciones sobre evaluación de docentes y estudiantes, b) con el propósito de identificar los factores que inciden en la variabilidad de concepciones, c) que profundizan en la relación entre las concepciones y las prácticas de evaluación y de los docentes y d) otras temáticas distintas. Veámoslos.

3.1. Análisis de los tipos de concepciones sobre evaluación del aprendizaje

Los primeros estudios en los que vamos a profundizar son aquellos que persiguen conocer las concepciones sobre evaluación. Entre ellas, cobra especial relevancia el trabajo realizado en la última década por Gavin Brown y sus colaboradores en Nueva Zelanda, el cual ha inspirado investigaciones posteriores con el mismo objeto de estudio (Brown, 2008; 2011; Brown y Harris, 2012; Brown y Hirschfeld, 2008; Brown, Irving, Peterson y Hirschfeld, 2009; Brown, Kennedy, Fok, Chan y Yu, 2009; Harris, Brown y Harnett, 2014; Hirschfeld y Brown, 2008).

En una serie de trabajos iniciados por Brown y su equipo se propusieron conocer las concepciones acerca de la evaluación de docentes de Educación Primaria y Secundaria. El enfoque de su investigación fue cuantitativo y el análisis de los datos se realizó a través de modelos de ecuaciones estructurales. Para la recogida de datos se utilizó el inventario "*Teachers' Conceptions of Assessment questionnaire (TCoA-III)*". Los hallazgos de su investigación confirman que existen cuatro grandes concepciones que los docentes tienen sobre la evaluación del aprendizaje del alumnado (Brown, 2003, 2004; Harris y Brown, 2008, 2009; Brown y Hirschfeld, 2008):

- *La evaluación como mejora*, comprendida como una actividad que proporciona información útil para transformar la práctica docente así como el aprendizaje de los estudiantes.
- *La evaluación como herramienta de rendición de cuentas de la escuela*, entendida como un instrumento para que los docentes demuestren que realizan correctamente su trabajo y que éste sirva para que los estudiantes alcancen los estándares educativos planteados y sea, por lo tanto, de calidad.
- *La evaluación como instrumento de rendición de cuentas del propio estudiante*, considerada como el momento en el que los estudiantes demuestran sus aprendizajes y el grado de consecución de los objetivos planteados.
- *La evaluación como un proceso irrelevante*, concibiendo la evaluación como un proceso poco útil para el aprendizaje ya que solo sirve para calificar y clasificar al alumnado, lo cual afecta a su autoconcepto y no ayuda a su aprendizaje.

Profundizando en cómo se relacionan estas cuatro concepciones halladas en sus estudios, Brown (2003, 2004, 2006, 2009) determina que los docentes que tienen una concepción cercana a la evaluación como mejora están también moderadamente de acuerdo con la creencia de que la evaluación es una herramienta de rendición de la escuela ($r_{xy} = 0,58$). Por otro lado, los profesores y profesoras que muestran menos cercanía con que la evaluación es irrelevante son los que muestran un acuerdo más elevado con la evaluación como mejora ($r_{xy} = -0,69$). Las concepciones de rendición de cuentas tanto de la escuela como de los propios estudiantes están moderadamente correlacionadas también ($r_{xy} = 0,58$). Estos resultados muestran que los docentes o estudiantes no se ubican únicamente en una concepción de evaluación sino que, aunque suelen mostrar más acuerdo con una de ellas, sus creencias también tienen elementos o características

del resto de concepciones. La investigación también revela que estas concepciones se mantienen constantes a lo largo del tiempo (Harris y Brown, 2008).

En el ámbito español, Ana Remesal (2011), de la Universidad de Barcelona, investigó las concepciones docentes de las funciones de la práctica evaluativa con un estudio cualitativo a través de entrevistas a treinta profesores. Los hallazgos principales del estudio son que las concepciones de evaluación se encuentran en un continuo con dos polos: un polo pedagógico y un poco social. Los docentes que se encuentran en el extremo de la concepción pedagógica de la evaluación consideran que ésta regula y ayuda al aprendizaje del alumnado mientras que los que se sitúan en el polo social defienden que la práctica evaluativa es un instrumento de acreditación social que sirve para clasificar y calificar a los estudiantes, no para aportarles información sobre cómo mejorar su aprendizaje. Frente a esta concepción dual de la evaluación, los resultados del estudio de Remesal (2011) reflejan que casi la mitad de los docentes tienen una concepción mixta más cercana al polo pedagógico, un 16% una creencia puramente pedagógica y otro 16% una concepción social. La autora concluye con que en general los docentes siguen mostrando concepciones arraigadas a las prácticas más tradicionales de evaluación.

Unos años más tarde, el investigador tunecino Sahbi Hidri (2015) en una investigación similar a la de Brown, se planteó conocer las concepciones de evaluación de los docentes de educación Secundaria y de Universidad. Utilizando el “*Teachers’ Conceptions of Assessment questionnaire* (TCoA-III)” y un cuestionario complementario adecuado al contexto de Túnez, Hibri (2015) utilizó el análisis factorial para analizar los datos. Los principales hallazgos son que el modelo de concepciones resultante estaba conformado por tres creencias de evaluación, concretamente que la evaluación sirve para la mejora, para la rendición de cuentas o que es irrelevante. Las concepciones de los docentes de mejora y rendición de cuentas correlacionaron positivamente. Estos hallazgos difieren en cierta medida por las investigaciones previas realizadas tanto por Brown como por otros investigadores. El autor tunecino concluye que estas diferencias pueden deberse a la escasa formación en evaluación que recibe el profesorado y la mala reputación de la evaluación en su país, pudiendo incidir en las concepciones del profesorado sobre las prácticas evaluativas.

Con el propósito de conocer las concepciones de evaluación de los profesores de Secundaria y relacionarlas con su práctica profesional, el profesor de la Universidad de Wellington Astuti Azis (2014, 2015) realizó un estudio mixto en el contexto de Indonesia. Teniendo en cuenta las concepciones encontradas en la investigación de Brown (2004, 2006, 2009), los participantes de este estudio consideran que la evaluación a) sirve para mejorar la enseñanza y el aprendizaje, b) sirve para la rendición de cuentas de la escuela y c) es útil para que los estudiantes demuestren aquello que han aprendido. Por otro lado, los resultados muestran que los docentes no están de acuerdo con que la evaluación sea irrelevante. La parte del estudio referente a la vinculación de las creencias con sus prácticas las mostramos en el epígrafe dedicado a este aspecto del presente artículo.

Otro aspecto en el que se centran estos estudios es no tan solo en las concepciones de docentes sino también en las de los estudiantes. Con este fin, Brown y Hirschfeld (2008) realizaron un estudio con 3.469 alumnos de educación secundaria con el objetivo de determinar las concepciones de los alumnos acerca de la evaluación. El análisis de

ecuaciones estructurales arrojó que los estudiantes tienen cuatro tipos diferentes de concepciones: a) evaluación como mejora el rendimiento y el aprendizaje, b) la evaluación hace que los estudiantes rindan cuenta y sean responsables, c) la evaluación es irrelevante, y d) la evaluación es agradable. Los autores concluyen que la primera concepción correlaciona positivamente con el rendimiento, mientras que los estudiantes con las concepciones c y d son los que peor desempeño obtienen.

Estos resultados nos permiten hallar similitudes entre las concepciones de docentes y estudiantes. Así, tres de las creencias son iguales para ambos mientras que hay una concepción diferente (para los docentes la evaluación como rendición de cuenta de la escuela frente a que la evaluación es agradable para el alumnado). En este sentido, existe coherencia y estabilidad también en las concepciones del alumnado, del mismo modo que sucede con las de los profesores y profesoras (Brown, 2003, 2004; Harris y Brown, 2008; Brown y Hirschfeld, 2008).

En un estudio posterior, Brown, Irving, Peterson y Hirschfeld (2009) matizaron que las concepciones de los estudiantes están más cercanas a la evaluación como una práctica formativa y no de control por parte del docente. Poco después, Brown y Harris (2012) complementaron dichos resultados llevando a cabo un nuevo estudio con el objetivo de conocer si existían diferencias en las concepciones de los estudiantes en función de su nivel de escolarización. Los resultados muestran que los estudiantes de educación secundaria, frente a los de primaria, puntúan más alto en concebir la evaluación como irrelevante, frente a concebirla como externa, como herramienta de mejora o como agradable. No obstante, en general, todos los estudiantes, sin diferencias significativas, conciben la evaluación en mayor medida como una herramienta que ayuda a mejorar su aprendizaje.

En Educación Secundaria, Peterson e Irving (2008) indagaron sobre las concepciones que tiene los estudiantes sobre la evaluación y su comunicación. El estudio, de corte cualitativo, llevó a cabo cinco grupos focales en cuatro escuelas de Nueva Zelanda. Los resultados señalan que los estudiantes conciben la evaluación como útil cuando ésta sirve para mejorar su aprendizaje e irrelevante cuando no proporciona información útil para ellos. Los autores concluyen que su estudio confirma de forma cualitativa los hallazgos de los estudios cuantitativos anteriores realizados por Brown y colaboradores.

En una línea similar de trabajo, Segers y Tillema (2011) realizaron una investigación con 712 estudiantes holandeses con el objetivo de conocer sus concepciones acerca de la finalidad que tiene la evaluación. Los análisis factoriales realizados apuntan a que existen cinco concepciones sobre la finalidad de la evaluación: a) es beneficiosa para el aprendizaje, b) sirve como instrumento para medir lo que los estudiantes han aprendido, c) resulta agradable a los estudiantes, d) mejora el clima de apoyo y colaboración en el aula, y e) sirve para la rendición de cuentas. Los autores señalan que los estudiantes que perciben la evaluación como un instrumento para la rendición de cuentas la asocian a una práctica profundamente injusta.

Conocidas las concepciones de profesorado y alumnado, otro aspecto investigado por Brown y sus colaboradores, es comparar las concepciones de evaluación en distintos países. Así, en distintos estudios realizados se compararon las concepciones de profesores y estudiantes de Nueva Zelanda con los de Australia, España, Irán, Chipre y China. Los resultados por países permiten afirmar que existen diferencias entre las creencias de los participantes neozelandeses y del resto de países. No obstante, las

concepciones de los docentes de los distintos países son un fiel reflejo del pensamiento de profesores y alumnos acerca de la naturaleza y el propósito de la evaluación, la enseñanza y los planes de estudio de dichos estados. Las diferencias explicativas se deben a la diversidad cultural y de los sistemas educativos y son coherentes con las políticas educativas de los distintos países analizados. Como pinceladas de los resultados de cada uno de los estudios:

- En el caso de Queensland (Australia), Brown y Lake (2006) se realizó un estudio *ex post facto* y un análisis a través ecuaciones estructurales. Los resultados determinaron que existen similitudes en las concepciones de evaluación como mejora e irrelevante pero diferencias en las creencias sobre la evaluación como rendición de cuentas. Por su contexto de gestión y de evaluaciones externas, los docentes de Nueva Zelanda tenían una visión más cercana a la rendición de cuentas de los estudiantes a través de la evaluación, frente a un menor acuerdo por los profesores y profesoras de Queensland. En este caso, ambos grupos de profesores también muestran múltiples concepciones que correlacionan entre sí de forma estable.
- En el contexto español, el propósito del estudio fue comparar las concepciones de evaluación de docentes en formación neozelandeses ($n=324$) y españoles ($n=672$). Los autores consideran que es importante conocer las concepciones de evaluación de los estudiantes de magisterio, dado que tienen una amplia experiencia como estudiantes y que los modelos utilizados para el análisis de las concepciones con los docentes en ejercicio no sirven para dichos estudiantes. En este caso, los hallazgos confirmaron que existen diferencias estadísticas significativas en las concepciones de los docentes en formación. Mientras que para los estudiantes de Nueva Zelanda, sus concepciones de la evaluación como mejora fueron inversas a la creencia de que la evaluación es irrelevante y positiva pero moderada con la rendición de cuentas, para los docentes en formación españoles las correlaciones fueron mucho más fuertes en ambos casos (Brown y Remesal, 2012). Como conclusión, los estudiantes de magisterio de España tienen una concepción sobre la evaluación más cercana a la mejora y la rendición de cuentas del propio estudiante.
- Con el propósito de conocer cómo conciben los profesores la evaluación en Chipre en comparación a Nueva Zelanda, Brown y Michaelides (2011) desarrolló un estudio *ex post facto* y posteriormente un análisis factorial exploratorio. Los resultados muestran que los docentes de Chipre tienen concepciones ligeramente diferentes a los docentes de Nueva Zelanda, debidas a las políticas y forma de llevarse a cabo la evaluación en ambos contextos. Las principales diferencias se encuentran en que para los profesores de Chipre las concepciones de la evaluación como mejora tienen una correlación moderada en considerar la evaluación como negativa o irrelevante y positiva con que la práctica evaluativa suponga una rendición de cuentas por partes de la escuela.
- Brown, Pishghadam y Sadafian (2014) analizaron si 760 estudiantes universitarios iraníes tienen una estructura de concepciones similar a los neozelandeses. Para ello, se hizo un análisis factorial confirmatorio y se obtuvo que las concepciones que tenían diferían ligeramente, siendo: a) la evaluación mejora el aprendizaje y enseñanza, b) la evaluación sirve para que las escuelas y

facultades rindan cuentas, y c) la evaluación tiene efectos negativos sobre el aprendizaje.

- En China, los profesores Gavin Brown y Lingbiao Gao (2015) examinaron cómo diversas investigaciones describen las concepciones y las prácticas de evaluación a través de un meta-análisis de ocho estudios de entrevistas y encuestas. En el análisis de las concepciones de evaluación de los docentes de China estos difieren al compararlos con los neozelandeses en considerar que la evaluación solo servirá para la mejora de la práctica docente y del aprendizaje del alumnado siempre y cuando cumpla su función ineludible de rendición de cuentas, tanto de la escuela como de los estudiantes.

Más allá de identificar las concepciones, otros investigadores buscaron comparar las creencias entre docentes y estudiantes universitarios (Fletcher, Meyer, Anderson, Johnston y Rees, 2012). Para ello, los autores realizaron un cuestionario para ambos grupos participantes y lo analizó con modelos de ecuaciones estructurales. Las diferencias en las concepciones de ambos grupos fueron claras. Mientras que los docentes universitarios concibieron la evaluación como un proceso que ayuda a mejorar la enseñanza y el aprendizaje, los estudiantes afirmaron que ésta solo sirve para que ellos rindan cuentas, siendo irrelevante para su aprendizaje. Para los autores es fundamental reflexionar acerca de la importancia de que la evaluación sea un proceso compartido entre docentes y estudiantes, sino, difícilmente podrá ayudar a los estudiantes y por ende, a la mejora del sistema educativo.

En síntesis, las investigaciones analizadas nos permiten afirmar que existen distintas concepciones de evaluación, que en todas las investigaciones se sintetizan en cinco: a) la evaluación como mejora, b) la evaluación como rendición de cuenta de la escuela, c) la evaluación como rendición de cuentas de los estudiantes, d) la evaluación es irrelevante, y e) la evaluación tiene efectos negativos sobre el aprendizaje. Las diferencias en las concepciones se originan en primer lugar por los participantes de la investigación, encontrando distintas concepciones si son estudiantes de educación obligatoria, futuros profesores o docentes en activo. En segundo lugar también encontramos diferencias por el nivel educativo, pudiendo estar las concepciones vinculadas a la evaluación que se desarrolla en Primaria, Secundaria o Educación Superior. En tercer lugar, las concepciones parecen ser estables en el tiempo y están relacionadas con el rendimiento en los estudiantes. Por último, las diferencias hemos visto que también pueden deberse al contexto del estudio, pudiendo afirmar que las concepciones se ven enormemente influidas por las políticas evaluativas del país o las directrices evaluativas del centro educativo. La influencia del contexto así como de otros factores se va a analizar en profundizar en el próximo apartado del presente artículo.

3.2. Estudios sobre los factores que inciden en la variabilidad de concepciones de la práctica evaluativa

Una segunda línea de estudio se centra en conocer los factores que inciden en las concepciones y su variabilidad. En uno de sus estudios primigenios, el profesor Gavin Brown (2004) afirmó que, inicialmente, las concepciones no se ven determinadas por el contexto en el que se desarrollan o por experiencia previa. Para ser más precisos, el autor concluyó que el número de años en la educación, el número de años de experiencia profesional y la situación socioeconómica de sus escuelas no influye en las concepciones que los docentes tenían en relación con evaluación. Frente a esta afirmación, otros

investigadores si han hallado diferencias en las concepciones de evaluación en función de distintos factores.

Efectivamente, el contexto es uno de los aspectos más estudiados por su relevancia en la conformación de las creencias de la evaluación del aprendizaje de los estudiantes (Rubie-Davies, Flint y McDonald, 2012; Vandeyar y Killen, 2007). La investigación ha evidenciado que las características sociales, culturales, económicas y políticas del país, ciudad o barrio donde trabajen los docentes o estudien los alumnos influyen directamente en las creencias implícitas que tengan de evaluación (Davis y Neitzel, 2010; James y Pedder, 2006; Prieto, 2008; Vandeyar y Killen, 2007).

Para la investigadora chilena Marcia Prieto (2008), las concepciones se ven influidas por el contexto en el que se desarrollan. Así, las creencias docentes están determinadas por el nivel socioeconómico, cultural y lingüístico de los estudiantes y del centro escolar. Los contenidos, los instrumentos y la finalidad de la evaluación planteada por el profesor cambian de una escuela a otra y de un estudiante u otro. Los profesores que trabajan en centros educativos ubicados en contextos desafiantes evaluarán teniendo en cuenta que sus estudiantes tienen una menor capacidad o están en desventaja social. En este sentido, la autora defiende que “la diversidad social y cultural que les caracterizan y por lo tanto, la evaluación se transforma en la vía que determina su trayectoria escolar, sobre la base de la calidad de su capital cultural” (Prieto, 2008, p. 134).

Vandeyar y Killen (2007), en una investigación de carácter etnográfico llevada a cabo en cuatro centros escolares de Sudáfrica, analizaron las creencias de sus docentes y concluyeron que el sistema dentro del cual los profesores y profesoras enseñan influye directamente en sus concepciones de evaluación. En este sentido, si el centro educativo hace hincapié en el contenido, las imposiciones políticas en términos evaluativos y las pruebas de evaluación externa, no es sorprendente que los profesores tengan concepciones de evaluación próximas a que la función de ésta es la rendición de cuentas del alumno o alumna y de la propia escuela. Por el contrario, si el centro educativo se preocupa del aprendizaje de sus estudiantes y de que la evaluación sea una herramienta de ayuda para estos, los docentes muestran concepciones más próximas a la mejora. En este sentido, la política, cultura y directrices evaluativas que sigan las escuelas determinan las concepciones de la práctica evaluativa de su profesorado.

No obstante, Vandeyar y Killen (2007) señalan que es muy iluso pensar que si cambian las políticas o cultura de evaluación de los escolares se transformaran automáticamente las concepciones de los docentes. La construcción de las creencias implícitas es un proceso complejo y por ello, el cambio en las concepciones también lo es. Ilustrando esta idea, el profesor británico Harry Torrance (1995) afirma que:

Creer que las concepciones de los docentes van a cambiar simplemente porque las políticas y los contextos escolares cambien es ingenuo, sin embargo, la capacidad y la voluntad de los docentes para implicarse con los cambios en el proceso de evaluación en particular, si se ve influida por sus ideas y creencias arraigadas. Por lo tanto, en un proceso de cambio evaluativo es fundamental implicar a los docentes en ella y que compartan su propósito, de lo contrario, no van a cambiar sus concepciones ni con ello su práctica evaluativa. (p. 55)

Aportando más datos empíricos de las ideas presentadas sobre la influencia del contexto en las concepciones de evaluación de los profesores, el estudio de Brown, Hui, Yu y Kennedy (2011) es un referente al respecto. En su investigación realizada en China los autores destacan la incidencia del contexto en la evaluación de los estudiantes,

considerando que la cultura, así como el estatus social de los docentes incide en la forma en que éstos conciben la enseñanza y la evaluación. Así, estos autores determinan que mientras que algunos factores pueden ser estables en toda la población, otras, como las concepciones de evaluación, pueden ser legítimamente diferente por las diferencias sociales y culturales existentes en las escuelas y en las características sociales de estas.

En este sentido, la investigación fenomenográfica de Murillo, Hidalgo y Flores (2016) con docentes de educación primaria y secundaria determinó que el contexto donde estuviese ubicada la escuela donde el docente desarrolla su labor influye en el tipo de concepciones de evaluación que tenían. Así, los profesores que trabajan en contextos desafiantes muestran unas concepciones de evaluación más cercanas a la inclusión, adaptación e individualización de la evaluación, mientras que los docentes de escuelas situadas en entornos favorables tienen creencias más cercanas a la excelencia, la objetividad y la igualdad en la evaluación. En palabras de los propios autores:

El contexto determina las concepciones y con ello las prácticas. De esta forma, frente al supuesto tratamiento indiferenciado por parte de las Administraciones de todos los centros, la investigación muestra que la segregación escolar configura enseñanzas muy diferentes en función del contexto socio-económico. Si tenemos bien asumida la idea de que hay que dar un mayor apoyo al estudiante que más lo necesita, no hay razones para no extrapolar esa idea a los centros. [...] Al final, es el contexto el que está determinando el tipo de evaluación, de enseñanza y con ello el futuro de los niños y las niñas.

El segundo factor analizado que puede incidir en las concepciones docentes de evaluación es las expectativas que los docentes tienen hacia sus alumnos y alumnas. Aunque no existen muchas investigaciones empíricas al respecto, queremos destacar el estudio ya mencionado realizado por Saloshna Vandeyar y Roy Killen (2007) de la Universidades de Pretoria y Newcastle respectivamente. En su estudio sobre la influencia del contexto, los autores matizan que dicho entorno y cultura del centro educativo determina las ideas implícitas que tienen sobre sus alumnos. Estas ideas previas influyen en las expectativas que cada profesor tiene de sus estudiantes, determinando el nivel de la enseñanza, los contenidos y la evaluación. Como conclusión de la investigación, los autores consideran que el entorno influye en lo que el docente espera de sus alumnos, pudiendo esperar o pensar que los estudiantes que se encuentran en centros educativos en entornos favorables pueden aprender más por estar en dicho contexto y las expectativas de los profesores serán mayores hacia estos estudiantes.

En tercer lugar, diversos estudios manifiestan que otro aspecto que influye en las concepciones de evaluación son las experiencias previas vividas con el proceso de enseñanza-aprendizaje. Así, tal como se ha vivido la evaluación como estudiantes determina las concepciones de los propios alumnos y de los profesores. Los estudios de Smith, Hill, Cowie y Gilmore (2014) y DiLoreto (2013) aportan datos al respecto. El primer de ellos ha sido realizado con estudiantes de magisterio de Nueva Zelanda. Los resultados muestran que sus concepciones próximas a la evaluación sumativa hicieron que les resultase complicado ser evaluados a través de pruebas abiertas y cualitativas. Esas concepciones tan arraigadas a las pruebas finales y objetivas se deben, en voz de sus propios protagonistas, a cómo han sido evaluados siendo alumnos y especialmente a cómo han sentido, vivido y gestionado la evaluación a lo largo de su escolaridad.

En esta línea, la afirmación “las concepciones son contextuales y vivenciales” (p. 2) cobra relevancia. La investigadora americana Melanie Anne DiLoreto adaptó en 2013 el

inventario de Brown de *Conceptions of Assessment (TCOA-III)* para conocer las concepciones de evaluación de los estudiantes de magisterio y de formación superior. Su análisis defiende que la evaluación debe servir para la mejora del propio aprendizaje así como para rendir cuentas de los conocimientos adquiridos. DiLoreto (2013) concluye que los participantes mostraron una fuerte relación entre dar cuentas de lo aprendido y mejorar, fruto de experiencias de evaluación pasadas.

Como aprendizaje de este conjunto de investigaciones, fruto de su construcción social, las concepciones no están libres del entorno en el que crecen, se relacionan y trabajan docentes o estudiantes, sino más bien al contrario. Asimismo, las vivencias sociales y educativas vividas por los estudiantes y las expectativas del profesorado también determinan sus concepciones de evaluación, y con ello, su aprendizaje y desarrollo integral.

3.3. La brecha entre las concepciones de evaluación y la práctica evaluativa

En tercera instancia, la relación entre las concepciones sobre evaluación y la propia práctica es un tema recurrente en las investigaciones centradas en conocer las ideas implícitas de docentes. La investigación ha demostrado que existe una evidente influencia de las creencias en la práctica evaluativa del docente. En efecto, la forma en que los docentes conciben la evaluación tiene repercusiones directas en las prácticas evaluativas del docente y éstas en el aprendizaje de los estudiantes (Prieto, 2008). Si el propósito de profundizar en las concepciones es impactar en la práctica de evaluación de los profesores, cobra relevancia conocer si estas ideas preconcebidas y la forma en que evalúan en sus clases son coherentes entre ellas.

Mostrando esta preocupación, Vandeyar y Killen (2007) analizaron las concepciones y sus prácticas de evaluación de docentes de escuelas sudafricanas de educación primaria a partir de tres estudios de caso con un enfoque etnográfico. El análisis de las entrevistas, las observaciones, los diarios docentes y la lectura de documentos clave demostraron la clara relación entre las concepciones sobre evaluación del profesorado y su práctica. Los autores demostraron que diferentes concepciones conducían a diferentes prácticas de evaluación. En este sentido, se encontraron las siguientes relaciones entre las concepciones y la práctica evaluativa de los profesores:

- Los docentes con una concepción de la evaluación como una herramienta útil para recopilar información para la toma de decisiones, la utilizaban como parte integrante de la enseñanza y el aprendizaje.
- Los educadores que ven la evaluación como algo que podría usarse para hacer que los estudiantes rindan cuentas por su propio aprendizaje, favorecerían la evaluación formal, sumativa y de alto rendimiento.
- Los profesores que conciben la evaluación como necesaria pero no importante para su aprendizaje favorecerían las prácticas de evaluación sumativa y cuasi formativa que ayudasen a generar una calificación detallada para rendir cuentas a la escuela.
- Los docentes que consideran que la evaluación es en gran medida irrelevante evitaban la evaluación formativa.

Las prácticas de evaluación observadas eran altamente consistentes con las concepciones de los docentes sobre la evaluación. En este sentido, las concepciones previas de

evaluación de los y las profesoras conducen a la alteración de sus métodos de evaluación. El estudio concluye con que los cambios educativos impuestos no promueven un cambio real en la prácticas evaluativas de los docentes. Para que ese cambio sea factible es necesario que el profesorado sea consciente de sus concepciones implícitas y que reciban formación y apoyo para un cambio de su propia práctica.

En coherencia con los resultados de Vandeyar y Killen (2007), Astuti Azis (2014, 2015) llevó a cabo un estudio con un enfoque mixto destinado a conocer las concepciones y prácticas de evaluación. Los resultados mostraron que aunque los docentes querían que la evaluación mejorase el aprendizaje de los estudiantes su concepción entraba en muchas ocasiones en conflicto con las evaluaciones externas impuestas por las políticas del país, siendo difícil cambiar dichas prácticas. Esta distancia entre las ideas implícitas y la práctica confirma que determinados factores externos en muchas ocasiones pueden interferir en que su evaluación refleje sus concepciones más arraigadas (Azis, 2015).

Ilustrando también la brecha entre concepciones y prácticas de evaluación, James y Pedder (2006) realizaron una investigación cuantitativa con 558 profesores en Inglaterra con el objetivo de conocer las concepciones sobre evaluación de los docentes y si estas son coherentes con su práctica profesional con el “*Staff Questionnaire (SQ)*”. Los resultados muestran que existe una clara brecha entre las prácticas y las creencias implícitas sobre evaluación del profesorado. Por lo tanto, a pesar de que estos profesores concebían la evaluación como una herramienta para facilitar el aprendizaje explícito y promover la autonomía del aprendizaje, sus prácticas evaluativas únicamente se orientaban a medir el rendimiento de sus estudiantes. Los autores concluyen que esta disparidad entre las concepciones y las prácticas se debe a la influencia y presiones que el contexto británico imprime en las prácticas de los profesores ya que tienen un claro carácter de rendición de cuentas de rendimiento de los estudiantes (Barnes et al., 2015).

En esta misma línea, Oscarson y Apelgren (2010) plantearon un estudio ex post facto con 605 docentes de educación secundaria que enseñaban idiomas de distintas lenguas. Con 20 de los docentes de la muestra se realizó a posteriori un estudio cualitativo para profundizar en dichas concepciones. Los resultados muestran que las evaluaciones se centran en el rendimiento del alumnado, por su concepción de rendición de cuentas. Los programas nacionales suecos en términos de dominio de la lengua y lenguas extranjeras como el inglés han impregnado profundamente las ideas implícitas y las prácticas evaluativas de los docentes, quienes se centran en observar la comunicación oral de su alumnado y medir sus aprendizajes. Los autores concluyen que encontraron ciertos conflictos conceptuales entre, por un lado, la evaluación, los exámenes y la calificación realizada por los docentes en sus aulas y, por otro lado, la interpretación lógica de las directrices y los planes de estudios que establecen los objetivos nacionales, generando problemas y una brecha entre las concepciones de evaluación y su práctica real.

Otros estudios se han focalizado en las concepciones de estudiantes de magisterio y las prácticas de evaluación que reciben como estudiantes universitarios o que han recibido en su vida escolar previa (Eren, 2010; Smith, Hill, Cowie y Gilmore, 2014; Yang y Tsai, 2009). En general, estos estudios determinan que sus experiencias previas y el contexto educativo influyen en sus concepciones. A continuación detallamos cada uno de estas investigaciones y sus principales hallazgos.

Por un lado, Yang y Tsai (2009) realizaron una investigación fenomenográfica con 163 estudiantes universitarios taiwaneses, con el propósito de conocer sus concepciones de

evaluación y compararlas con la forma en que eran evaluados en sus clases a través de la evaluación por pares. Los datos reflejan que existen dos concepciones claras entre los estudiantes: una relacionada con una evaluación que provoca el aprendizaje superficial y fragmentando y otra vinculada con que la evaluación favorece que el aprendizaje sea cohesivo y profundo. Los estudiantes que mostraban una concepción cercana a la segunda de las dos concepciones presentadas progresaron más en la evaluación por pares planteada por los docentes. Por otro lado, los alumnos con concepciones de que la evaluación no ayuda al aprendizaje no sacaron provecho de la evaluación entre compañeros ni influyó en sus compañeros. Los investigadores demuestran, una vez más, que nuestras ideas implícitas influyen en nuestras prácticas, y en este caso, en la evaluación por pares planteada en el aula universitaria.

Un año más tarde, Eren (2010) comparó las concepciones de 304 futuros docentes de primaria y secundaria y las prácticas de evaluación con un estudio de encuesta con el “*Staff Questionnaire (SQ)*” desarrollado originalmente por James y Pedder (2006). Los resultados globales del estudio mostraron que existen dos concepciones principales: una evaluación tradicional frente a una evaluación constructivista. La mayoría de los docentes en formación se mostraron cercanos a la concepción que relaciona la evaluación como un proceso constructivista del aprendizaje del estudiante y que ayuda a redirigir la enseñanza del profesorado. Asimismo, los autores concluyeron que los docentes llevaban a cabo una evaluación más tradicional que constructivista, a pesar de tener concepciones próximas al constructivismo.

En tercer lugar, Smith, Hill, Cowie y Gilmore (2014) profundizaron en las concepciones de los estudiantes de magisterio y cómo estas concepciones pueden impactar en su futura práctica docente. Para ello se llevó a cabo un estudio cuantitativo-cualitativo con 720 estudiantes de distintas universidades ubicadas en distintas regiones de Nueva Zelanda. Los principales hallazgos del estudio son que es necesario trabajar para cambiar las concepciones sobre evaluación del profesorado si el fin de la misma es ayudar a los estudiantes en su aprendizaje. Los estudiantes de magisterio señalan la importancia de la evaluación en el desarrollo del alumnado y conocer las concepciones es básico para elaborar planes de formación docente centrados en una práctica evaluativa formativa, cualitativa y compartida con los alumnos y alumnas.

El análisis de estos estudios nos aporta datos relevantes que nos permiten afirmar que las prácticas de evaluación del profesorado no son siempre un fiel reflejo de sus concepciones. Así, sus ideas implícitas en ciertas ocasiones están vinculadas con una evaluación entendida como reguladora y que mejora el aprendizaje, a través de un proceso cualitativo y formativo. No obstante, factores como el contexto donde se encuentran trabajando los docentes, sus experiencias previas vividas como estudiantes con la evaluación o las directrices de evaluación que tiene el centro pueden marcar irremediamente su práctica de evaluación. Cuando estos aspectos influyen en la evaluación del profesorado, esta se aleja de sus concepciones y convierten la práctica evaluativa en un proceso mucho más sumativo y dirigido a la rendición de cuentas.

3.4. Otros estudios actuales sobre concepciones de evaluación

En cuarto lugar, consideramos interesantes presentar aquellos estudios que se centran en estudiar otros aspectos relacionados con las concepciones sobre evaluación. Entre estos aspectos estudiados destacamos el estrés laboral docente, el uso de los datos de evaluación y los efectos de las evaluaciones estandarizadas.

La evaluación es una de las prácticas docentes que puede provocar situaciones de estrés por parte de los profesores y profesoras por su implicación en la promoción de sus estudiantes. Con esta idea de partida, Pishghadam, Adamson, Sadafian y Kan (2014) profundizaron en la relación entre las concepciones de evaluación y el *burnout* experimentando por los docentes. Con este fin, los investigadores desarrollaron un estudio cuantitativo con dos cuestionarios: el *"Teachers' Conceptions of Assessment questionnaire (TCoA-III)"* de Brown y colaboradores y el *"Maslach Burnout Inventory"*. Los participantes fueron 113 docentes de lengua inglesa iraníes. El análisis de regresión múltiple muestra una relación significativa entre las concepciones de evaluación y las tres dimensiones del *burnout* (agotamiento emocional, despersonalización y baja realización personal). Los docentes más cercanos a la concepción que considera que la evaluación es irrelevante son los que muestran menos estrés, mientras que los que consideran que la evaluación ayuda a la mejora del aprendizaje del alumnado son los que muestran un nivel de agotamiento más elevado. En este sentido, el profesorado más implicado en la evaluación de sus estudiantes era el que mostraba signos más elevados de *burnout*, habiendo una relación significativa positiva entre ambos.

En relación con la evaluación estandarizada, Atta Gebril y Gavin Brown (2013) analizaron los efectos de la evaluación externa en las concepciones de 301 estudiantes del último año de magisterio y 202 docentes en activo de escuela y universidad en Egipto. Los datos se recolectaron a través del *"Teachers' Conceptions of Assessment questionnaire (TCoA-III)"* y se utilizó un análisis factorial confirmatorio para extraer los resultados. El análisis de los datos refleja que los docentes y estudiantes egipcios muestran tres concepciones distintas de evaluación: la evaluación mejora el aprendizaje, la evaluación sirve para la rendición de cuenta de la escuela y la evaluación irrelevante. En la concepción de la evaluación como mejora, los docentes en formación puntúan más alto que los profesores en servicio. El modelo de concepciones refleja la imposición social de las evaluaciones externas, donde los profesores consideran que una escuela de alta calidad mejora el rendimiento de su alumnado para su rendición de cuentas.

Buscando la relación entre las concepciones y el uso de los datos resultado del proceso de evaluación, Gannon-Slater (2014) llevó a cabo un estudio cualitativo donde se analizó el uso de estos datos para la toma de decisiones como una estrategia de innovación para mejorar enseñanza y el aprendizaje. Los participantes fueron 16 docentes de Educación Secundaria. Los resultados mostraron que la mayoría de los maestros a menudo utilizan múltiples instrumentos para extraer suficientes datos para evaluar a sus estudiantes. En cuanto al uso de esos datos de evaluación, los docentes concebían que los datos obtenidos de evaluaciones estandarizadas no servían para cambiar o mejorar su enseñanza. Los autores del estudio concluyen que las normas institucionales, organizacionales e incluso profesionales influyen directamente en el uso que los docentes dan a distintos tipos de datos así como a las concepciones de los mismos. Las ideas implícitas de los profesores son las que dan sentido al uso de dichos datos, viéndose influidos por las normativas de evaluación existentes en cada contexto.

Estudios centrados en determinar la relación de las concepciones sobre evaluación y distintos aspectos confirman que las ideas implícitas del profesorado no influyen únicamente en su práctica evaluativa sino también en su nivel de estrés sufrido en el aula y el uso de los datos de la evaluación. Asimismo, las evaluaciones externas existentes en cada país también influyen en sus concepciones sobre la práctica evaluativa, otorgándole

cada vez más peso a la rendición de cuentas frente a la mejora. Estas investigaciones nos permiten afirmar que las concepciones intervienen en la práctica profesional docente, y no tan sólo en cómo evalúan.

4. Reflexiones a modo de cierre

Actualmente se están multiplicando las evaluaciones estandarizadas externas en todo el mundo, y en consecuencia de ello, la presión que reciben docentes y estudiantes. Frente a esta realidad, crece el interés por diseñar sistemas de evaluación válidos y fiables, pero también que sean útiles y justos. Las concepciones de evaluación son un elemento clave en esta situación, dado que están profundamente vinculadas con la práctica evaluativa. Actualmente, existen varias investigaciones realizadas destacan la importancia del tema como hemos querido presentar en este artículo, pero aún nos queda mucho por saber.

En síntesis, podemos afirmar la investigación sobre las concepciones de los profesores y profesoras sobre la evaluación es fundamental para comprender la complejidad del proceso de enseñanza-aprendizaje (especialmente de la práctica evaluativa). A pesar de que numerosos investigadores son conscientes de la importancia de la relación entre la evaluación, las concepciones y la práctica, aún sigue siendo un campo de estudio en el que se ha profundizado insuficientemente (Dayal y Lingam, 2015). De acuerdo con Dana Opre (2015):

La realización de estudios ambiciosos que pretendiesen analizar la naturaleza e intensidad de esta relación entre las concepciones y la evaluación, podría aportar una contribución esencial al desarrollo de este campo de investigación. Además, el resultado de Este tipo de esfuerzo podría abrir nuevas direcciones de investigación o incluso un nuevo enfoque de los objetivos de la investigación. (p. 232)

La revisión de las investigaciones relacionadas con las concepciones sobre la evaluación de los estudiantes nos ha permitido extraer varios aprendizajes. En primer lugar, las concepciones están profundamente influidas por el contexto donde se encuentren los docentes y estudiantes, marcando claramente su práctica evaluativa. Otro aspecto que determina las concepciones son las experiencias pasadas vividas como estudiantes de evaluación, sucediendo que, en muchos casos, los docentes reproducen aquello que han vivido en el pasado.

En segundo lugar, también sabemos que las concepciones sobre evaluación no siempre se traducen en una práctica evaluativa concreta. Así, concebir la evaluación como mejora de la evaluación no supone llevar a cabo una práctica evaluativa con el mismo fin, especialmente porque las políticas educativas, las directrices del centro educativo y las evaluaciones externas dirigen la evaluación desarrollada en el aula. Esto se traduce en una clara brecha entre concepciones previas y práctica evaluativa, siendo fundamental ir estrechando esa diferencia.

En efecto, otro aspecto aprendido en esta revisión de investigaciones es que en función de cómo los profesores conciben la evaluación, esto impactará en su estrés laboral vivido en el aula, en cómo utilicen los datos de la evaluación de sus estudiantes y en cómo viven las evaluaciones estandarizadas externas. Esto supone tener en cuenta que es importante conocer las ideas implícitas del profesorado y estudiantes acerca de la evaluación ya que

determina varios aspectos de su práctica profesional, lo que es fundamental conocer si queremos mejorar la evaluación.

Los avances realizados en esta línea de investigación nos ofrecen muchas respuestas pero también nos plantean nuevas preguntas. Entre estos interrogantes destacamos la necesidad de avanzar un paso más en la relación entre concepciones y práctica evaluativa, no tan solo señalando su brecha contextual sino viendo que nexos existen entre ambos aspectos. Asimismo, es interesante profundizar en cómo las vivencias de estudiantes determinan las concepciones de evaluación de los docentes, y como hacerlas explícitas para transformar su práctica.

Otra posible línea de investigación futura que se vislumbra de esta revisión es profundizar en el conocimiento y la comprensión del origen de las diferentes concepciones sobre evaluación de estudiantes, docentes y futuros docentes. La investigación ha evidenciado que tienen ideas implícitas distintas, pero aún no se ha avanzado en qué causa dicha diferencia en sus concepciones. Otro aspecto interesante en el estudio de las concepciones es profundizar si las características socio-demográficas personales tanto de docentes como estudiantes influyen en sus concepciones, pudiendo, de este modo, conocer mejor los factores de su variabilidad.

Este artículo ha permitido aproximarnos a las principales investigaciones relacionadas con las concepciones sobre evaluación. Las aportaciones revisadas son una inspiración para seguir profundizando en qué tipos de concepciones existen, cuáles son los factores que influyen y determinan las creencias previas o cuáles son las causas de la brecha entre estas y la práctica. Los distintos estudios analizados han puesto de manifiesto la relevancia de profundizar en las concepciones si queremos comprender mejor la práctica y transformarla para lograr una evaluación de estudiantes de mayor calidad, más equitativa y más justa.

Referencias

- Atjonen, P. (2014). Teachers' views of their assessment practice. *Curriculum Journal*, 25(2), 238-259. doi:10.1080/09585176.2013.874952
- Atkinson, T. y Claxton, G. (2000). *The intuitive practitioner: On the value of not always knowing what one is doing*. Buckingham: Open University Press.
- Azis, A. (2012). Teachers' conceptions and use of assessment in student learning. *Indonesian Journal of Applied Linguistics*, 2(1), 40-52. doi:10.17509/ijal.v2i1.72
- Azis, A. (2014). *Indonesian junior secondary school teachers' conceptions of assessment: A mixed methods study*. Tesis Doctoral. Victoria Universidad de Wellington, Nueva Zelanda.
- Azis, A. (2015). Conceptions and practices of assessment: a case of teachers representing improvement conception. *Teflin Journal*, 26(2), 129-154. doi:10.15639/teflinjournal.v26i2/129-154
- Barnes, N., Fives, H. y Dacey, C. M. (2015). Teachers' beliefs about assessment. En H. Fives y M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 284-300). Londres: Routledge.
- Brown, G. T. L. (2003). *Teachers' conceptions of assessment*. Tesis Doctoral. Universidad de Auckland, Nueva Zelanda.

- Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301-318. doi:10.1080/0969594042000304609
- Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged version 1, 2. *Psychological Reports*, 99(1), 166-170. doi:10.2466/pr0.99.1.166-170
- Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. Auckland: Nova Science Publishers.
- Brown, G. T. L. (2011). Self-regulation of assessment beliefs and attitudes: A review of the Students' Conceptions of Assessment inventory. *Educational Psychology*, 31(6), 731-748. doi:10.1080/01443410.2011.599836
- Brown, G. T. L. y Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education*, 2, 1-19.
- Brown, G. T. L. y Harris, L. (2012). Student conceptions of assessment by level of schooling: further evidence for ecological rationality in belief systems. *Australian Journal of Educational & Developmental Psychology*, 12, 46-59.
- Brown, G. T. L. y Hirschfeld, G. H. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3-17. doi:10.1080/09695940701876003
- Brown, G. T. L. y Lake, R. (noviembre, 2006). Queensland teachers' conceptions of teaching, learning, curriculum and assessment: Comparisons with New Zealand teachers. Comunicación presentada en la *Australian Association for Research in Education Conference*, Adelaide, Australia.
- Brown, G. T. L. y Michaelides, M. (2011). Ecological rationality in teachers' conceptions of assessment across samples from Cyprus and New Zealand. *European Journal of Psychological Education*, 26(3), 319-337. doi:10.1007/s10212-010-0052-3
- Brown, G. T. L., Lake, R. y Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education*, 27(1), 210-220. doi:10.1016/j.tate.2010.08.003
- Brown, G. T. L., Pishghadam, R. y Sadafian, S. S. (2014). Iranian university students' conceptions of assessment: Using assessment to self-improve. *Assessment Matters*, 6, 5-33.
- Brown, G. T., Hui, S. K., Flora, W. M. y Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5), 307-320. Ddoi:10.1016/j.ijer.2011.10.003
- Brown, G. T. L., Irving, S. E., Peterson, E. R. y Hirschfeld, G. H. (2009). Use of interactive-informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning and Instruction*, 19(2), 97-111. doi:10.1016/j.learninstruc.2008.02.003
- Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S. y Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347-363. doi:10.1080/09695940903319737
- Chen, J. A., Morris, D. B. y Mansour, N. (2015). Science teachers' beliefs: Perceptions of efficacy and the nature of scientific knowledge and knowing. En H. Fives y M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 370-386). Londres: Routledge.

- Coll, C. y Remesal, A. (2009). Mathematics teachers' conceptions about the functions of assessment in compulsory education. *Journal for the Study of Education and Development*, 32(3), 391-404. doi:10.1174/021037009788964187
- Davis, D. S. y Neitzel, C. (2010). The relationship between students' reading orientations and their strategic activity during a collaborative reading task. *Reading Psychology*, 31(6), 546-579. doi:10.1080/02702710903256593
- Dayal, H. C. y Lingam, G. I. (2015). Fijian teachers' conceptions of assessment. *Australian Journal of Teacher Education*, 40(8), 42-58. doi:/10.14221/ajte.2015v40n8.3
- DiLoreto, M. A. (2013). Multi-group invariance of the conceptions of assessment scale among university faculty and students. Tesis doctoral. Universidad de Southern Mississippi, Mississippi.
- Ekeblad, E. y Bond, C. (noviembre, 1994). The nature of a conception: Questions of context. Comunicación presentada en el congreso *Phenomenography: Philosophy and Practice, Proceedings*. Universidad de Queensland, Australia.
- Eren, A. (2010). Consonance and dissonance between Turkish prospective teachers' values and practices: conceptions about teaching, learning, and assessment. *Australian Journal of Teacher Education*, 35(3), 27-48. doi:10.14221/ajte.2010v35n3.3
- Fives, H. y Gill, M. G. (2014). *International handbook of research on teachers' beliefs*. Nueva York, NY: Routledge.
- Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P. y Rees, M. (2012). Faculty and students conceptions of assessment in higher education. *Higher Education*, 64(1), 119-133. doi:10.1007/s10734-011-9484-1
- Gannon-Slater, N. (2014). *Reframing teachers' conceptions of assessment and data use*. Tesis Doctoral. Universidad de Illinois, Chicago.
- Gebriel, A. y Brown, G. T. (2014). The effect of high-stakes examination systems on teacher beliefs: Egyptian teachers' conceptions of assessment. *Assessment in Education: Principles, Policy and Practice*, 21(1), 16-33. doi:10.1080/0969594X.2013.831030
- Griffiths, T., Gore, J., y Ladwig, J. (noviembre, 2006). Teachers' fundamental beliefs, commitment to reform, and the quality of pedagogy. Comunicación presentada en la Conferencia anual *Proceedings Australian Association for Research in Education*. Universidad de Adelaide, Australia.
- Harris, L. R. (2008). A phenomenographic investigation of teacher conceptions of student engagement in learning. *The Australian Educational Researcher*, 35(1), 57-79. doi:10.1007/BF03216875
- Harris, L. R. y Brown, G. T. L. (octubre, 2008). *New Zealand teachers' conceptions of the purpose of assessment: Phenomenographic analyses of teachers' thinking*. Comunicación presentada en la *Australian Association for Research in Education Conference*, Brisbane, Australia.
- Harris, L. R. y Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practices*, 16(3), 365-381. doi:10.1080/09695940903319745
- Harris, L. R., Brown, G. T. L., y Harnett, J. A. (2014). Understanding classroom feedback practices: A study of New Zealand student experiences, perceptions, and emotional responses. *Educational Assessment, Evaluation and Accountability*, 26(2), 107-133. doi:10.1007/s11092-013-9187-5

- Hasselgren, B. y Beach, D. (1997). Phenomenography-a "good-for-nothing brother" of phenomenology? Outline of an analysis. *Higher Education Research & Development*, 16(2), 191-202.
- Hidri, S. (2016). Conceptions of assessment: Investigating what assessment means to secondary and university teachers. *Arab Journal of Applied Linguistics*, 1(1), 19-43.
- Hirschfeld, G. H. y Brown, G. T. (2009). Students' conceptions of assessment: Factorial and structural invariance of the SCoA across sex, age, and ethnicity. *European Journal of Psychological Assessment*, 25(1), 30-38. doi:10.1027/1015-5759.25.1.3
- James, M. y Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109-138. doi:10.1080/09585170600792712
- Marshall, B. y Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133-149. doi:10.1080/02671520600615638
- Martín, E., Mateos, M., Martínez, P., Cervi, J., Pecharromán, R. y Villalón, R. (2006). Las concepciones de los profesores de educación primaria sobre la enseñanza y el aprendizaje. En J. I. Pozo, N. Scheuer, M. Pérez Echeverría, M. Mateos, E. Martín y M. De la Cruz (Eds.), *Nuevas formas de pensar la enseñanza y el aprendizaje. Las concepciones de profesores y alumnos* (pp. 171-187). Barcelona: Graó.
- Marton, F. (1981). Phenomenography – Describing conceptions of the world around us. *Instructional Science*, 10, 177-200. doi:10.1007/BF00132516
- Murillo, F. J. e Hidalgo, N. (2015). Enfoques fundamentantes de la evaluación de estudiantes para la justicia social. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 43-61.
- Murillo, J., Hidalgo, N. y Flores, S. (2016). Incidencia del contexto socio-económico en las concepciones docentes sobre evaluación. *Profesorado. Revista de Currículum y Formación del Profesorado*, 20(3).
- Murillo, F. J., Martínez-Garrido, C. e Hidalgo, N. (2014). Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España. *Estudios sobre Educación*, 27, 91-113.
- Opre, D. (2010). *Expertiza didáctica universitară*. Praga: ASCR.
- Opre, D. (2015). Teachers' Conceptions of Assessment. *Procedia - Social and Behavioral Sciences* 209, 229-233. doi:10.1016/j.sbspro.2015.11.222
- Oscarson, M. y Apelgren, B. M. (2011). Mapping language teachers' conceptions of student assessment procedures in relation to grading: A two-stage empirical inquiry. *System*, 39(1), 2-16. doi:10.1016/j.system.2011.01.014
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332. doi:10.3102/00346543062003307
- Peterson, E. R. e Irving, S. E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction*, 18(3), 238-250. doi:10.1016/j.learninstruc.2007.05.001
- Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. En F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257-315). Charlotte, NC: Information Age Publishing.
- Pishghadam, R., Adamson, B., Sadafian, S. S., y Kan, F. L. (2014). Conceptions of assessment and teacher burnout. *Assessment in Education: Principles, Policy & Practice*, 21(1), 34-51. doi:10.1080/0969594X.2013.817382

- Pope, N., Green, S. K., Johnson, R. L. y Mitchell, M. (2009). Examining teacher ethical dilemmas in classroom assessment. *Teaching and Teacher Education*, 25(5), 778-782. doi:10.1016/j.tate.2008.11.013
- Pozo, J. (2006). La cultura del aprendizaje en la sociedad del conocimiento. En J. Pozo, N. Scheuer, M. P. Pérez Echevarría, M. Mateos, E. Martín y M. De la Cruz (Eds.), *Nuevas formas de pensar la enseñanza y el aprendizaje. Las concepciones de profesores y alumnos* (pp. 29-53). Barcelona: Graó.
- Pozo, J., Scheuer, N., Mateos, M. y Pérez Echeverría, M. P. (2006). Las teorías implícitas sobre la enseñanza el aprendizaje. En J. Pozo, N. Scheuer, M. P. Pérez Echevarría, M. Mateos, E. Martín y M. De la Cruz (Eds.), *Nuevas formas de pensar la enseñanza y el aprendizaje. Las concepciones de profesores y alumnos* (pp. 95-134). Barcelona: Graó.
- Pratt, D. D. (1992). Conceptions of teaching. *Adult education quarterly*, 42(4), 203-220.
- Prieto, M. (2008). Creencias de los profesores sobre evaluación y efectos incidentales. *Revista de Pedagogía*, 29(84), 123-144.
- Prieto, M. y Contreras, G. (2008). Las concepciones que orientan las prácticas evaluativas de los profesores: un problema a develar. *Estudios Pedagógicos*, 34(2), 245-262.
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472-482. doi:10.1016/j.tate.2010.09.017
- Rubie-Davies, C. (2015). Teachers' instructional beliefs and the classroom climate. En H. Fives y M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 266-283). Nueva York, NY: Routledge.
- Rubie-Davies, C. M., Flint, A. y McDonald, L. G. (2012). Teacher beliefs, teacher characteristics, and school contextual factors: What are the relationships?. *British Journal of Educational Psychology*, 82(2), 270-288. doi:10.1111/j.2044-8279.2011.02025.x
- Säljö, R. (1994). Minding action: Conceiving of the world versus participating in cultural practices. *Nordisk Pedagogik*, 2, 71-80.
- Sambell, K., McDowell, L. y Brown, B. (1997). 'But is it fair?': An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, 23(4), 349-371. doi:10.1016/S0191-491X(97)86215-3
- Scott, S., Webber, C., Lupart, J., Aitken, N. y Scott, D. (2014). Fair and equitable assessment practices for all students. *Assessment in Education: Principles, Policy & Practice*, 21(1), 52-70. doi:10.1080/0969594X.2013.776943
- Segers, M. y Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment?. *Studies in Educational Evaluation*, 37(1), 49-54. doi:10.1016/j.stueduc.2011.03.008
- Smith, L. F., Hill, M. F., Cowie, B. y Gilmore, A. (2014). Preparing teachers to use the enabling power of assessment. En C. Wyatt-Smith, V. Klenowski y P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 303-323). Dordrecht: Springer.
- Thompson, A. G. (1992). *Teachers' beliefs and conceptions: A synthesis of the research*. Nueva York, NY: Macmillan Publishing.
- Torrance, H. (1995). Teacher involvement in new approaches to assessment. En H. Torrance (Ed.), *Evaluating authentic assessment: problems and possibilities in new approaches to assessment* (pp. 44-56). Philadelphia, PA: Open University Press.
- Van den Berg, B. (2002). Teachers' meanings regarding educational practice. *Review of Educational Research*, 72, 577-625. doi:10.3102/00346543072004577

- Vandeyar, S. y Killen, R. (2007). Educators' conceptions and practice of classroom assessment in post-apartheid South Africa. *South African Journal of Education*, 27(1), 101-115.
- Winterbottom, A., Bekker, H. L., Conner, M. y Mooney, A. (2008). Does narrative information bias individual's decision making? A systematic review. *Social science & medicine*, 67(12), 2079-2088. doi:10.1016/j.socscimed.2008.09.037
- Yang, Y. F. y Tsai, C. C. (2010). Conceptions of and approaches to learning through online peer assessment. *Learning and Instruction*, 20(1), 72-83. doi:10.1016/j.learninstruc.2009.01.003

Breve CV de los autores

Nina Hidalgo Farran

Doctorando en Educación en el área de Métodos de Investigación y Diagnóstico en Educación en la Universidad Autónoma de Madrid, con una beca de doctorado de carácter competitivo otorgada por el Ministerio. Es diplomada en Ciencias de la Educación en la especialidad de Maestra de Educación Infantil por la Universidad de Lleida y licenciada en Psicopedagogía por la Universidad Autónoma de Madrid. Es miembro del grupo de investigación Cambio Educativo para la Justicia Social (GICE) y miembro del equipo editorial de la Revista Internacional de Educación para la Justicia Social. Asimismo es editora de la Revista de la Revista Iberoamericana de Evaluación Educativa. Cuenta con diversas publicaciones, como artículos, reseñas y capítulos de libro. Página personal: <http://www.gice-uam.es/Nina.Hidalgo>. ORCID ID: 0000-0002-0056-8566. Email: nina.hidalgo@uam.es.

F. Javier Murillo Torrecilla

Profesor titular del Área de Métodos de Investigación y Diagnóstico en Educación de la Universidad Autónoma de Madrid. Coordinador del grupo de investigación Cambio Educativo para la Justicia Social (GICE). Fue Coordinador General del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), de la UNESCO, y Director de Estudios del Centro de Investigación y Documentación Educativa (CIDE) del Ministerio de Educación de España. Es Coordinador de la Red Iberoamericana de Investigación sobre Cambio y Eficacia Escolar (RINACE), Director de REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación, de la Revista Iberoamericana de Evaluación Educativa y de la Revista Internacional de Educación para la Justicia Social. Página de docencia: <http://www.uam.es/javier.murillo>. ORCID ID: 0000-0002-8003-4133. Email: javier.murillo@uam.es.

Volumen 10

Núm. 2

A1.3. Publicación 3.
Incidencia del contexto socio-económico en las
concepciones docentes sobre evaluación

revista de curriculum y
formación del profesorado





Incidencia del contexto socio-económico en las concepciones docentes sobre evaluación

Impact of socio-economic context in the teachers' conceptions of assessment

F. Javier Murillo; Nina Hidalgo y Sofía Flores

Universidad Autónoma de Madrid

nina.hidalgo@uam.es

RESUMEN¹

Las concepciones que tienen los docentes sobre la evaluación tienen una fuerte incidencia sobre sus prácticas evaluativas y, con ello, sobre el desarrollo de los estudiantes. Por ello, existe una fructífera línea de investigación que busca conocer esas concepciones y cómo inciden en el proceso de enseñanza y aprendizaje. Por otro lado, hay evidencias empíricas de que las actuaciones de los docentes se definen en gran medida por el contexto socio-económico del centro donde el docente trabaja. Sin embargo, se sabe aún poco de cómo el contexto socio-económico del centro incide en

¹ Este estudio está incluido en la Convocatoria de Proyectos de Investigación I+D+i "Escuelas para la Justicia Social" (EDU2011-29114), dirigido por F. Javier Murillo (Universidad Autónoma de Madrid)

esas concepciones docentes. El presente estudio busca conocer las concepciones de evaluación que tienen los profesores de educación Primaria y Secundaria, y determinar si existen diferencias en función del contexto socio-económico del centro en el que trabaja. Para ello se ha desarrollado un estudio cualitativo de carácter fenomenográfico. Los participantes diez docentes de Educación Primaria y Secundaria, la mitad que trabajan en un contexto socio-económico favorable y la otra mitad en un contexto desafiante. Para obtener los datos se utilizó la entrevista fenomenográfica. Los resultados muestran importantes diferencias en las concepciones sobre los referentes de evaluación, sobre los instrumentos usados y su adaptación, y sobre la devolución de la información. Lo más relevante es la diferente concepción que tienen de lo que es evaluación justa. Para los docentes de entornos favorables una evaluación justa está ligada a la objetividad y a la transparencia de las evaluaciones, los instrumentos y los criterios de corrección; mientras que para los docentes que trabajan en contextos desafiantes está más relacionada con la adaptación y la medición integral de los avances del estudiante. La conclusión es la importancia de tener en cuenta el contexto socio-económico, tanto para desarrollar investigaciones como para poner en marcha políticas de apoyo a los centros.

PALABRAS CLAVE: Concepciones de evaluación, Evaluación de estudiantes, Contexto del centro educativo, Investigación fenomenográfica, Evaluación Justa.

ABSTRACT

The teachers' conceptions of assessment have a strong impact in their evaluative practices, and with this, in the students' development. Thus, exist a fruitful research line that seeks to understand these conceptions and how they affect the process of teaching and learning. Furthermore, there is empirical evidence that the actions of teachers are defined greatly by the socio-economic context of the school

where the teacher works. However, still little is known of how the socio-economic context of the school commit in teachers' conceptions. The current study looks for know the assessment conceptions of Primary and Secondary teachers, in order to determine whether there are differences depending on the socio-economic context of the school where teachers work. For this, we have developed a phenomenographic approach with 10 teachers, half working in a favorable socio-economic context and the other half in a challenging context. Phenomenographic interviews were used to obtain data. The results show significant differences in assessment conceptions relating to the evaluation model, the instruments and methods used and their adaptation means, and the return of assessment information. The most relevant is the different conception they have of what is fair assessment. For teachers, which work in favorable environments, a fair assessment is linked to the objectivity and transparency of assessments, tools and correctioncriteria; while for teachers which work in challenging contexts is more related to adaptation and comprehensive measure of student progress. The conclusion is the importance of having in mind the socio-economic school context, both to develop research as well as to implement policies to support schools.

KEYWORDS: Evaluation conceptions, Assessment, School context, Phenomenographic research, Fair evaluation.

1. INTRODUCCIÓN

En la actualidad se están multiplicando las investigaciones que buscan indagar en las concepciones de los docentes hacia la practica evaluativa (p.ej., Coll y Remesal, 2009; Griffiths, Gore y Ladwig, 2006; Harris y Brown, 2009; Pope, Green, Johnson y Mitchell, 2009; Remesal, 2011). La idea es sencilla, dado que las concepciones son el primer paso para la acción, sólo cambiando las concepciones

podemos mejorar las prácticas y, con ello, el desarrollo de los estudiantes (Murillo, Martínez-Garrido e Hidalgo, 2014).

Efectivamente, todo ser humano, por naturaleza, construye ideas, pensamientos y juicios sobre y a partir de sus propias experiencias. Esto lo lleva a concebir el mundo de una forma diferente de los otros, y que incide, de una forma u otra en su modo de actuar en los diferentes ámbitos de la vida. Está claro que la causa de una acción no es necesariamente es reflejo de que así piensa o concibe el mundo una persona, pero sin duda lo determina de una forma importante. Por ejemplo, que un profesor evalúe de cierta forma no necesariamente refleja que crea que sea la más adecuada, pero sin duda se ve afectada por esas ideas implícitas. En cualquier caso, la manera en que el docente concibe a la evaluación impacta en su propia práctica evaluativa y, por tanto, en el proceso de enseñanza-aprendizaje (Marshall y Drummond, 2006).

Por otra parte, existe una buena cantidad de evidencias que indican que el contexto socio-económico en el que está la escuela determina la forma de actuar de los docentes. Así, resulta de especial interés la investigación de Lupton (2005) donde estudió cómo influye el contexto social en la dinámica del centro. El planteamiento era sencillo, seleccionó las escuelas de los barrios más pobres de Inglaterra y analizó cómo el contexto marcaba el trabajo de docentes y directivos. Y se dio cuenta de que los profesores no podían desarrollar su tarea de forma normal, aunque quisieran; sabían la teoría y se creían buenos profesores, pero no podían mantener un buen clima, organizar lecciones estructuradas, proponer actividades desafiantes o atender a la diversidad..., si tenían que mantener unas rutinas de supervivencia tales como el orden en el aula.

Enlazando ambas ideas, nos planteamos la siguiente pregunta de investigación: ¿Incide el contexto socio-económico de la escuela en las concepciones de los docentes sobre la práctica evaluativa?

2. REVISIÓN DE LA LITERATURA

El profesor estadounidense M. Frank Pajares (1992) afirmó en un influyente artículo que prestar atención a las concepciones del profesorado debería de ser un tema prioritario para la investigación educativa. Si las concepciones determinan el comportamiento de los docentes, solo conociéndolas es posible modificarlas para mejorar la enseñanza y el desarrollo de los estudiantes.

Las concepciones se definen como un sistema organizado de creencias, originadas en las experiencias del individuo y desarrolladas a través de las interacciones en las que participa (Remesal, 2011). Es importante señalar que, a pesar de la apariencia individualista, las concepciones tienen un carácter social, ya que se fundamentan e interrelacionan en fenómenos culturalmente compartidos (Van den Berg, 2002). Dado que estas creencias se desarrollan a través de las experiencias personales de los individuos, los investigadores concluyen que las concepciones son omnipresentes, e influyen y son influidas por las interacciones del individuo con el mundo. Así, una concepción es la una construcción mental sobre aquello que un sujeto interpreta, vive y entiende una realidad, pero en ningún caso como lo experimenta ya que el interés de las concepciones radica en conocer que piensa un sujeto sobre un proceso determinado (Marton, 2000).

Las concepciones docentes serían, por tanto, un marco o estructura de creencias a través del cual los docentes interpretan e interactúan día a día en el aula (Pratt, 1992; Thompson, 1992); De esta forma, las concepciones de los profesores son

influidas por el contexto en el que se desarrolla el profesor, y definen su práctica docente. Queda claro que dicha influencia no es siempre coherente o explícita, pero sin duda las concepciones son un factor decisivo que determinan el proceso de enseñanza-aprendizaje (Griffiths, Gore y Ladwig, 2006).

Prieto y Contreras (2008) insisten en precisar que los profesores no construyen sus creencias de manera intencionada y consciente, por el contrario, las han internalizado inconscientemente en sus estructuras cognitivas y emocionales, a partir de una gran diversidad de experiencias personales y escolares. Las experiencias escolares acaecidas en el contexto en el que se desempeñan les han facilitado la construcción social de creencias respecto de la educación, la enseñanza, el aprendizaje o la evaluación.

De esta forma, no es extraño que el estudio de las concepciones de los docentes se ha convertido en una fructífera línea de investigación educativa (Fives y Gill, 2014). Dentro de ella, tiene especial relevancia el estudio de las concepciones docentes sobre sus prácticas evaluativas (Atjonen, 2014; Azis, 2012; Brown, 2003, 2004, Brown y Hirschfeld, 2008; Brown y Remesal, 2012; Brown, Lake y Matters, 2011; Chen, Brown, Hattie y Millward, 2012; Harris y Brown, 2008; Winterbottom et al., 2008; entre otros).

La influencia de las creencias en la práctica evaluativa del docente es clara: la forma en que los docentes conciben la evaluación tiene repercusiones directas en las prácticas evaluativas del docente y éstas en el aprendizaje de los estudiantes. Efectivamente, toda evaluación llevada a cabo por un docente ejerce un poder determinando tanto en las prácticas pedagógicas como en el aprendizaje de los estudiantes (Prieto, 2008; Prieto y Contreras, 2008).

En los últimos años, y especialmente a partir de los trabajos de Brown (p.ej., Brown, 2003, Brown, Lake y Matters, 2011; Brown y Hirschfeld, 2008; Harris y Brown, 2008), es posible reconocer cuatro grandes concepciones que tienen los docentes de la evaluación y su repercusión en la práctica evaluativa (Azis, 2012):

- *La evaluación como mejora.* Los docentes que conciben la evaluación desde esta perspectiva consideran que la información que surge de la evaluación sirve para cambiar y mejorar su propia enseñanza así como el aprendizaje de los estudiantes. Estos docentes utilizan diversas estrategias para evaluar a los estudiantes, aportándoles información cuantitativa y cualitativa tanto para los propios docentes como para los estudiantes (Brown, 2003, 2004; Harris y Brown, 2008; Brown y Hirschfeld, 2008; Black, Harrison, Lee, Marshall y William, 2002).
- *La evaluación entendida como rendición de cuentas de la escuela.* Los docentes que evalúan en base a estas creencias consideran que la evaluación sirve para dar cuenta de su propia labor docente y de la consecución de estándares, así como de la aportación educativa de la escuela en la sociedad. Para estos docentes la evaluación es responsabilidad de la escuela y sirve para demostrar la calidad de su enseñanza.
- *La concepción de evaluación como rendición de cuentas del propio estudiante.* Esta concepción sostiene que los estudiantes son responsables de su propio proceso de aprendizaje, por lo que el docente se limita a acreditar y calificar su logro académico (Harris y Brown, 2008). Desde esta perspectiva, los docentes deben ser lo más “objetivos” posibles con el fin de calificar correctamente a sus estudiantes ya que esto tiene consecuencias tales como la graduación, la selección o la presentación de informes públicos.

- *La evaluación como un proceso irrelevante.* Los docentes que conciben el proceso evaluativo desde esta perspectiva rechazan la evaluación ya que consideran que es un elemento más del proceso de enseñanza. Consideran que tener que calificar o valorar a los estudiantes afecta a su autonomía y profesionalidad así como al aprendizaje de los estudiantes. Además consideran que los datos obtenidos de la evaluación son poco válidos y menos fiables (Brown, 2003).

De entre las investigaciones recientes destacamos el trabajo Remesal (2011), tanto por su calidad como por haber sido desarrollado en nuestro contexto. Esta investigadora desarrolló una ambiciosa investigación sobre las concepciones de evaluación de los docentes de educación Primaria y Secundaria en España. El propósito del estudio es explorarlas concepciones acerca de las funciones de la evaluación de los docentes. En el estudio, de carácter cualitativo, se realizaron entrevistas en profundidad a treinta docentes. Las concepciones de evaluación son analizadas en base a la posición de los docentes en un polo pedagógico o social. Los docentes con concepciones puramente pedagógicas consideran la evaluación como un elemento regulador del aprendizaje de los estudiantes y de su propia práctica; las concepciones puramente sociales conciben la evaluación como un elemento acreditativo social, la cual no debe cambiar el proceso de enseñanza, considerando la enseñanza como algo puramente objetivo y rechazando los métodos cualitativos de evaluación. Los resultados del estudio muestran que un 40% de los docentes conciben la evaluación de forma mixta social (es decir, no puramente social per más cercano a éste polo que al pedagógico); un 24% tienen una concepción mixta pedagógica; un 16% con una concepción puramente pedagógica y otro 16% con una concepción puramente social. Estos datos muestran que los docentes siguen teniendo

concepciones tradicionales arraigadas de la evaluación del aprendizaje de los estudiantes.

Aun sabiendo que el contexto es un insumo fundamental para la elaboración de las concepciones de los docentes (p.ej., Brown, 2003), no hay mucha investigación que nos permita saber qué contextos generan qué concepciones (p.ej., Rubie-Davies, Flint y McDonald, 2012). Sin embargo, sobresalen las aportaciones de Marcia Prieto (2008). Esta autora chilena considera que las creencias están profundamente relacionadas con el contexto en el que se desarrollan, por lo que el nivel socioeconómico, cultural y lingüístico de los estudiantes y del centro influirá directamente en las concepciones y creencias de evaluación del docente. Según ella los contenidos y la forma de evaluarlos se determina en función del contexto del centro educativo. Desde esta perspectiva, los docentes situados en contextos más desafiantes o con dificultades deberán adecuarse más a las diferencias entre los estudiantes, “como la diversidad social y cultural que les caracterizan y por lo tanto, la evaluación se transforma en la vía que determina su trayectoria escolar, sobre la base de la calidad de su capital cultural” (Prieto, 2008, p. 134). Así, las creencias sobre los estudiantes que provienen de sectores más vulnerables se relacionan con su “menor capacidad”, su “poca dotación” o su inevitable “desventaja sociocultural y lingüística”.

Con todo ello, los objetivos de esta investigación son profundizar en las concepciones de evaluación de docentes españoles, y determinar si el contexto socio-económico del centro educativo en el que trabajan determina diferencias en las concepciones.

3. MÉTODO

3.1. ENFOQUE METODOLÓGICO

Se ha optado por usar la fenomenografía como forma de aproximación a las concepciones de los docentes (Bouden y Walsh, 2000; Marton, 1986; Richardson, 1999). Marton (1994) describió a la fenomenografía como un enfoque con base empírica que procura identificar el número limitado de maneras cualitativamente distintas de como diferentes personas vivencian, conceptualizan, perciben y entienden diversos tipos de fenómenos. Se caracteriza por ser un método que cambia continuamente, ya que es construido en conjunto por aquellos que participan en ella. Su objeto son las diferentes (o semejantes) maneras de cómo somos conscientes del mundo, cómo experimentamos, comprendemos, aprendemos y vemos los diversos fenómenos y situaciones del mundo. Desde una perspectiva fenomenográfica, el interés del investigador no se centra en los procesos en sí mismos sino en cómo los sujetos vivencian y perciben dichos procesos (Reed, 2006).

3.2. PARTICIPANTES

Una cuestión fundamental en relación con los estudios fenomenográficos es la elección de los informantes claves de la investigación, tanto el número como la variabilidad. Los participantes en esta investigación son diez profesores de diferentes características que trabajan en centros de la Comunidad de Madrid. Su selección se hizo mediante un muestreo no probabilístico por cuotas con tres criterios de selección:

- *Contexto socio-económico del centro educativo:* favorable o desafiante. Los indicadores que nos han permitido seleccionar estos centros han sido: el nivel socioeconómico del barrio donde se encuentra el centro y de las familias de los estudiantes que asisten a ese centro, y las características sociales, culturales y lingüísticas de dichos estudiantes. Entre los centros favorables se encuentran

centros situados en los barrios madrileños de Chamartín, Valdezarza o Vicálvaro; y entre los centros desafiantes barrios como Villa de Vallecas, Orcasitas o Carabanchel.

- *Nivel educativo del centro educativo*, diferenciado entre los docentes de Educación Primaria y los de Educación Secundaria.
- *Titularidad del centro*, centros educativos públicos, concertados y privados.

En la tabla 1 se detalla la información, donde puede verse que existe una distribución equilibrada de dichos docentes.

Tabla 1. *Características de los participantes.*

DOCENTE	NOMBRE *	NIVEL EDUCATIVO	TITULARIDAD DEL CENTRO	CONTEXTO	ASIGNATURA	AÑOS DE EXPERIENCIA
1	Carolina	Primaria	Público	Desafiante	Tutor	10
2	Matías	Primaria	Público	Desafiante	Matemáticas	14
3	Lourdes	Primaria	Concertado	Desafiante	Tutor	3
4	Ana	Primaria	Concertado	Desafiante	Tutor	5
5	Alberto	Primaria	Concertado	Favorable	Tutor	8
6	Marta	Primaria	Concertado	Desafiante	Tutor	12
7	Pedro	Secundaria	Privado	Favorable	Humanidades	10
8	Juan	Secundaria	Concertado	Favorable	Lengua e Historia	25
9	Susana	Secundaria	Público	Desafiante	Historia	25
10	Laura	Secundaria	Privado	Favorable	Francés y Lengua	8

Nota: (*) Los nombres de los docentes que aparecen en la tabla de participantes son ficticios, con el fin de preservar su confidencialidad. Fuente: Elaboración propia.

3.3. INSTRUMENTO DE RECOGIDA DE INFORMACIÓN Y ANÁLISIS DE DATOS

El instrumento utilizado en el presente estudio es la entrevista fenomenográfica (Trigwell, 2000). Este método de obtención de información se basa en el auto-relato (Marton y Booth, 1997). La clave para llevar a cabo una entrevista fenomenográfica es permitir al participante reflexionar sobre su experiencia acerca de un fenómeno

(Bowden, 1996). Las experiencias captadas por la entrevista se "constituyen conjuntamente por el entrevistador y del entrevistado" (Marton, 1994, p.4427), entendido como una 'conversación entre dos compañeros sobre un tema de interés mutuo' (Kvale, 1996, p. 125).

La entrevista de tipo fenomenográfico es de carácter semi-estructurada con sólo unas pocas preguntas clave predeterminadas (Bowden, 1996; Marton, 1994; Reed, 2006). Esto no significa que la entrevista fenomenográfica no tenga foco ya que el objeto de estudio es el centro de enfoque del entrevistador en todo momento y guía la entrevista en todo momento.

Para dar respuesta a esta necesidad de una entrevista semi-estructurada, las entrevistas partieron de tres preguntas generadoras: "¿cómo evalúas?, ¿Cómo concibes tu evaluación? y ¿cómo sería tu evaluación ideal? Se recogieron datos individualmente, se grabaron y posteriormente se transcribieron íntegramente. El trabajo de campo fue desarrollado entre los meses de mayo a julio de 2014.

Acorde con el método fenomenográfico se ha seguido un criterio de proximidad al discurso de los participantes. El investigador trata de ser neutral a las ideas de los participantes en el estudio, analizando los datos obtenidos con la mente abierta y sin participación alguna de sus perspectivas (Orgill, 2002). Se registró anotando la presencia de conceptualizaciones pertenecientes a determinadas categorías y sub-categorías en el discurso de los docentes, separando los docentes que trabajan en centros de contextos favorables y desafiantes. La búsqueda de categorías llevo a encontrar un elemento que agrupa y caracteriza las diferentes concepciones de los docentes y es el concepto que tiene de evaluación justa. De esta forma se utilizó para aglutinar las diferencias en sus concepciones.

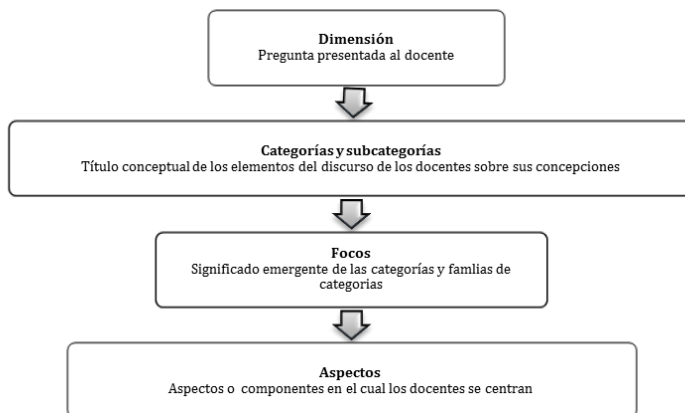


Figura 1: *Pasos seguidos en el análisis del discurso de los docentes.*

Fuente: Elaboración propia.

4. RESULTADOS

Los hallazgos encontrados en el presente estudio muestran una visión compleja y llena de matices que nos permite comprender las concepciones de los docentes sobre su práctica evaluativa. De las entrevistas fenomenográficas realizadas a diez docentes de la Comunidad de Madrid que trabajan en centros educativos situados en contextos favorables y en entornos desafiantes, y de su análisis, aparecen una serie de diferencias en las concepciones que tienen de su propia práctica y de una evaluación ideal. Estas diferencias generan el concepto de evaluación justa como elemento caracterizador y aglutinador de las mismas.

De esta forma se organizarán los resultados en dos bloques, las concepciones sobre su práctica evaluativa y la concepción que tienen de lo que es una evaluación justa.

4.1. CONCEPCIONES SOBRE SU PROPIA PRÁCTICA EVALUATIVA

¿Los docentes de centros educativos situados en contextos socio-económicamente favorables tienen concepciones sobre su propia práctica evaluativa diferentes de los docentes de centros en contextos desafiantes? Esta es la pregunta de investigación que orienta este apartado. Y la respuesta a la misma es un claro sí, pero con infinitud de matices.

Un primer elemento llamativo es el relativo a fuentes en las que se inspiran los docentes para hacer las pruebas. En las escuelas en contextos favorables es muy habitual usar **referentes externos**, esencialmente el libro de texto como base mínima y, a partir de ahí, subir más el nivel con los trabajado en el aula. Veámoslo con tres ejemplos de tres profesores de centros diferentes:

Para elaborar la prueba me baso en el libro de texto; bueno ellos tienen libro de texto, en nuestro caso tenemos dos libros de texto de la misma editorial [...]Les pregunto un poco de eso que lo que hemos visto en clase. [p4FAV:2]

Lo tenía preparado parte del año pasado, pero es que nunca doy todo igual, es decir, ciertas cosas tengo que cambiar porque a lo mejor cierto tipo de información pues he visto que no se lo he dado tan bien o he visto que ellos no lo han podido captar tan bien y entonces no insisto tanto en esto [...]. Por eso, además de basarme en el libro de texto, ahora hay material buenísimo para los profesores. Tenemos material de ampliación, tenemos material de refuerzo, y eso lo uso para los exámenes.[p8FAV:2]

Hay preguntas que tanto se han trabajado en clase con el libro como otras preguntas que han surgido en clase pues por ejemplo, de este control que era una hoja que no sale en el tema del libro, pero sí que trabajamos en clase.[p5FAV:2]

Los y las docentes de centros en contextos desafiantes, sin embargo, sus fuentes son más internas, más autorreferenciales que externas. En sus discursos aparecen constantemente la elaboración mediante el trabajo en equipo y la colaboración con otros docentes de ese curso para determinar los contenidos o destrezas mínimas hasta donde quieren que lleguen sus estudiantes; curiosamente ninguno nombra el libro de texto. Como comentan Carolina y Lourdes:

Yo coordino las pruebas de evaluación con mi compi del otro segundo y elaboramos juntas todo. O sea y entre otras cosas los exámenes. [p1DES:3]

Los exámenes los suelo elaborar yo y en muchas ocasiones en equipo, en colaboración, y suelo hacer exámenes muy variados y muy diferentes uno de otro. [p3DES:1]

Aunque la práctica totalidad de los docentes afirma usar una amplia variedad de **instrumentos de obtención de información** para la evaluación (escalas de actitudes, hojas de observación, diarios de clase...), el examen escrito, con preguntas de diferentes tipos sigue siendo tanto la prueba que más afirman usar los docentes, especialmente en Secundaria. Aun con esa pauta general, sí se observan algunas diferencias en función del contexto favorable-desafiante donde se sitúan los centros educativos.

Efectivamente, las informaciones apuntan a que los docentes que trabajan en centros cuyo contexto es favorable utilizan con mayor frecuencia y con más peso en la nota final pruebas cuantitativas. La concepción que tienen es que este tipo de pruebas son más objetivas y posibilitan tener información más clara a los estudiantes para mejorar su aprendizaje. Ilustrando esta idea, Juan y Laura afirman:

Utilizo dos exámenes por trimestre. Otra prueba de evaluación es la revisión del cuaderno de trabajo de clase, pero eso solo sirve para redondear la nota. El

peso de la nota lo tienen los exámenes escritos que realizan de los diferentes temas. [p9DES: 3]

Si tengo que ser sincera, es eso; los instrumentos de evaluación sobre todo son: exámenes como te estoy comentando, esas pruebas orales, pequeñas exposiciones o a veces les digo, como todavía no tienen capacidad de hablar, porque hasta tercero de la E.S.O. y son muy poquitas horas [...] y ahí les evalúo y les pongo nota de su participación y demás [...] pero lo que más cuenta es el examen. [p10FAV: 2]

Los docentes de escuelas en contextos desafiantes, aun usando también el examen escrito tradicional, muestran un mayor interés por el uso de pruebas "alternativas", que no muestren sólo la capacidad de los estudiantes de memorizar y redactar. Su preocupación es que deben también considerarse las diferentes habilidades y capacidades de sus estudiantes, el aprendizaje real de los estudiantes y el esfuerzo que realizan. Sirven tres ejemplos para ilustrar:

Aparte del control escrito, también evalúo la expresión oral, yo no hago en un día el control de la expresión oral, lo que pasa que sí que continuamente las actividades de clase un día le toca exponer a unos, otro día les toca exponer a otros, yo tengo mi hoja de observación que sin decir a ellos "yo te estoy evaluando" pero sí que yo le pongo cómo lo va haciendo, con la fecha y así sucesivamente otros días yo les voy preguntando [...] Otra cosa aparte es lo que apunto yo en mi diario de algo que me haya llamado más la atención [...] Los controles escritos cuentan en mi evaluación el 50% de la nota, el otro 50% son otro tipo de competencias que considero esenciales y que no pueden medirse con un examen. [p4DES:3]

Para mí lo más importante es el día a día de los chicos y su proceso de aprendizaje, que ellos sean conscientes. Por ello, el plan de trabajo tiene un 60 por ciento del porcentaje de la nota y los exámenes el otro 30 por ciento.

[p9DES:3]

...No suelo hacer exámenes como tal, lo veo de una manera de ponerle nota a un niño pero no se demuestran tanto sus habilidades.. [...] Procuro hacer ejercicios de evaluación en donde no se den cuenta de que están siendo evaluados y no como una prueba real, por lo que les motiva a realizarla de una manera mucho más adecuada...[p3DES: 3]

Muy relacionado con el anterior elemento es el hecho de que los profesores y profesoras en contextos desafiantes muestran una especial preocupación por **adaptar** las pruebas con situaciones con las que se puedan sentir identificados los estudiantes. Unas palabras que reflejan este planteamiento son las siguientes:

En una prueba, por ejemplo, redacté yo como una lectura de unos niños que eran de otro barrio y que iban a una escuela por la tarde y tal, que no eran ellos pero bueno para que ellos se sintieran así, más motivados y tal. Y claro, salió bastante bien, con eso se quedaron en seguida. [p4DES:2]

Los docentes en contextos desafiantes insisten en que la evaluación sea individualizada, adaptada a las necesidades y características de cada uno de los estudiantes. La ratio por aula, las evaluaciones externas, el tiempo del que disponen o la administración son algunos de los motivos señalados por los docentes que más dificultan lograr una verdadera evaluación individual y ajustada a cada estudiante. En palabras de los propios docentes:

Adaptar más las pruebas a cada alumno, porque realmente es muy difícil hacer un examen por ejemplo escrito, un control escrito para todos igual porque es

muy difícil. Tú vas a la media, al final a la media, a los objetivos medios pero hay alumnos que para ellos es muy fácil y para otros muy difícil, y al final yo hago el control y evidentemente para los alumnos de integración y los alumnos de compensatoria, no es el mismo control. [p2FAV: 7]

Para mí una evaluación ideal sería una evaluación individualizada, tener en cuenta todo lo que rodea al alumno, su implicación, su esfuerzo... como te decía, no sólo la prueba objetiva como es un examen. Es que cada niño es un mundo, cada niño, cada familia que hay detrás, entonces eso, valorar todo eso. [p1DES: 4]

Tener en cuenta las necesidades del alumno, sus capacidades, todo, hasta su entorno, tener en cuenta su situación familiar, todo. Yo creo que para evaluar a un alumno y sobretodo en una etapa de primaria y cuando son más pequeñitos, hay que tomar en cuenta no sólo los conocimientos que hayan adquirido que sí es súper importante pero también todo lo que hay a su alrededor. [p3DES: 4]

También se observa que, frente a los que imparten clase en centros favorecidos, los docentes de centros educativos ubicados en contextos desafiantes determinan como esencial enfocar la evaluación hacia un proceso más cualitativo. Los profesores consideran que las pruebas escritas no ofrecen una información real del aprendizaje de los estudiantes, y que es necesario conocer más instrumentos cualitativos de evaluación y llevarlos a cabo, para promover un proceso más continuo de seguimiento de los alumnos, pudiéndoles dar una retroalimentación más frecuente y centrada en las necesidades de cada estudiante. Así lo expresan Ana y Susana:

Y de alguna otra manera que a mí me gustaría que fuera más oral, que fuera más individual, necesitaría una hora para cada alumno [...] habría que hacer cada vez los exámenes más abiertos y no tan cerrados, no con respuestas tan

cerradas... pienso que las podría hacer un poco más abiertas... eso, que podría adaptarlas más al nivel de cada uno específico y bueno, el objetivo sería que no existieran las pruebas escritas. [p4DES: 5]

Y luego me parece más justo cuando hay diferentes tipos de preguntas o diferentes formas de demostrar lo que sabes... no me gustan los exámenes... sólo los exámenes, no me gustan... pero me gusta que... que los alumnos... aprendan a demostrar lo que saben poniéndolo en la práctica y dándoselo a los demás. [p9DES: 4]

Por último, los docentes de centros desafiantes consideran la necesidad de que la evaluación sea un proceso constante y que pueda repetirse, dando diferentes oportunidades a los estudiantes para demostrar su conocimiento y aprendizaje. La profesora Susana, lo expresa de la siguiente forma:

También es justo que se den oportunidades, que se dé más de una oportunidad porque... no puede haber una sola oportunidad eso no es justo, porque puedes tener mal día, porque puedes.... [p9DES: 3]

Ligada a esta idea de individualización (o no) de la evaluación, los docentes de zonas favorecidas asumen que la **devolución de los resultados** se realiza de forma grupal, así consideran que el proceso es más transparente. Alberto, profesor en un centro situado en un entorno favorable, indica que:

Yo les corrijo primero el examen, lo que pasa es que siempre quiero que lo corrijan y que hagan el intento de buscar qué es lo que hicieron mal y tal, entonces sí que dedico una clase para eso. [...]. Lo que suelo hacer es entregarles el control, primero lo corregimos de forma grupal para resolver las dudas. Realmente lo que les hago es, les entrego el control, en él viene la nota y un comentario sobre qué me ha parecido su control y cómo puede mejorar,

qué fallos ha tenido... y luego hago una corrección grupal [...]. En esa corrección grupal preguntan “¿y esto por qué está mal o...?”, sí, suelen preguntar cosillas, sí. [p5FAV:3]

Los docentes de contextos desafiantes, por su parte, se preocupan en que dicha devolución sea lo más individualizada posible. La idea que subyace es la necesidad de que los estudiantes pueden mejorar su aprendizaje con directrices claras, como expone la profesora Carolina:

... lo que les hago es, les entrego el control, en él viene la nota y un comentario sobre qué me ha parecido su control y cómo puede mejorar, qué fallos ha tenido... y luego hago una corrección grupal. [p1DES:2]

En esta misma línea, Ana, profesora de un centro concertado en contexto desafiante, nos comenta que:

Si puedo con todos lo comento uno a uno individualmente. Y si no, cada uno tiene señalado... no quiere decir mal, menos las tipo test, las otras yo no les tacho que está mal ni nada, por ejemplo si hay una falta de ortografía, si es una expresión que no está bien hecha o si es algo en que se han equivocado de que de eso no va la lectura, pues entonces les pongo como una flechita y ellos ya saben que ahí tienen que observar que algo pasa. Y entonces pues se comenta en voz alta, “pues yo ya me he dado cuenta de lo que es”. [p4DES: 4]

4.2. ¿QUÉ ES UNA EVALUACIÓN "JUSTA"?

De lo anteriores elementos es posible configurar a la concepción que tiene los docentes de qué es una evaluación "justa" para los estudiantes. Mientras que para los docentes de centros favorables ese concepto está ligado a la imparcialidad, a la idea de una evaluación objetiva, igual para todos. Para los de contextos desafiantes está

más relacionado con la idea de valorar el esfuerzo de cada estudiante; de ahí también la importancia de la adaptación individual de las evaluaciones.

Efectivamente, para los de contextos favorables la mejor prueba es aquella que no tiene margen de error. Es decir, que es absolutamente cerrada y que no admite discusión o debate. En esta línea, el profesor Juan, nos indica que:

Yo lo que busco es la objetividad [...]. Corrijo lo más objetivamente posible. Mira, la de verdadero y falso es muy fácil, 25 por 0,4. [...]Me gusta de tal manera que esté estructurado y que ellos sepan... en general no suelo tener protestas como “¿y por qué me pones aquí un 1,25 sobre 2?” y le digo “a ver, ¿cuántas causas había?” y efectivamente... la ortografía y la expresión, eso va aparte. Corrijo las 25, o las preguntas que haya, si alguna pregunta vale menos porque es más sueltita, aparece también en el examen. [p8FAV:3]

La objetividad y la transparencia son dos criterios que debe tener una evaluación justa para los docentes en contexto favorecidos:

Yo... muchas veces... yo utilizaría más actividades complementarias y utilizaría más otro tipo de trabajo... pero prima la objetividad. Te lo pide el sistema e incluso el propio alumnado. [p8FAV:4]

Creo que un componente importante de la justicia es la honestidad, entonces yo creo que tienen que haber una claridad en la evaluación y saber a qué atenerte; tener unos criterios establecidos y en función de esos criterios la persona que es evaluada sabe que va a ser evaluada y sabe para qué es la evaluación. [p4FAV:3]

[...] Hay un problema que son los objetivos... tú te tienes que ceñir a unos objetivos que te marca la ley, y se supone que si el niño no llega a esos

objetivos pues no puede pasar del cinco, se supone... no puede pasar de cinco en esa evaluación [p9FAV:5]

Una evaluación en la que se tengan en cuenta todo el entorno del alumno, todas sus características, su evolución, el proceso que ha tenido... [p5FAV:4]

Los profesores y las profesoras que enseñan en escuelas situadas en entornos desafiantes, por su parte, dan más importancia a la evolución del estudiante, al avance real de los estudiantes entre prueba y prueba de evaluación:

No corrijo de forma cerrada. Entonces lo que yo intento es valorar, es que ha habido un aprendizaje, que ha habido un avance y... el que haya una mitad de pruebas cortas... es para intentar que aprueben porque la segunda parte es mucho más difícil para los alumnos... [p9DES: 4]

Lo importante no es la nota que saquen en el examen escrito ni si avanzan puestos en el cálculo mental, sino saber de dónde a partido y hasta dónde ha llegado. No puede ser igual de un niño que parte con una nota de un 8 y llega a un 8,5 que otro que parte con un 3 y llega a un 7. En el primer caso parece que tu aportación ha sido muy poca, mientras que en el segundo caso ha de reconocerse el logro conseguido. La nota que saques de un examen no me dice mucho, no me preocupa, sólo me detecta las necesidades. Lo que yo intento valorar a lo largo del año es el esfuerzo y el avance que ellos han tenido desde su punto de partida hasta su punto final. [p2DES: 3]

Asimismo, un aspecto que destacan los docentes de escuelas en contextos desafiantes es ser capaces de evaluar verdaderamente el esfuerzo de los estudiantes. Muchos docentes indican la dificultad de saber cuánto se esfuerza un estudiante, pero inciden en la importancia de premiar a aquellos que más empeño ponen en la tarea. Así, los docentes exponen que:

En una evaluación justa sería necesario valorar mucho más a una persona que le costó desde el primero momento adquirir ese conocimiento y él se lo ha luchado, se lo ha peleado y ha conseguido adquirir ese conocimiento... o sea para mí no puede tener la misma nota el niño que me ha sacado las mismas cinco operaciones que desde el primer momento las coge fenomenal, que al niño que le ha costado un mes llegar a ellas. Porque a lo mejor al que le ha salido desde el primer momento fenomenal, podía haber llegado un escalón más arriba... [p6DES:4]

Siempre hay que seguir aprendiendo. Estoy aprendiendo a evaluar los esfuerzos, más que la nota en sí. Darle importancia al trabajo diario, al esfuerzo diario, y ver de dónde parten y hasta dónde llegan". [p2DES: 3]

Mira una evaluación ideal sería aquella evaluación que reflejase el esfuerzo del niño durante esos 15, 20 días o un año... si tú quieres evaluar contenidos concretos o una asignatura, que seas capaz de valorar el esfuerzo que el chico ha realizado.[p7DES: 4]

La que mide realmente el esfuerzo y si ese esfuerzo ha llevado ha llevado a conseguir los objetivos que he planteado, tomando en cuenta de dónde parte el alumno y a dónde ha llegado, sobre todo teniendo en cuenta de dónde parte. Porque no es la cuestión llegar al mismo punto sino ver, o sea, la que evalúa la cantidad de camino que has hecho tú relacionado con tu esfuerzo.[p4DES: 5].

¿Qué es una evaluación justa? ¿La que es objetiva, transparente e igualitaria, o la que es objetiva, diferenciadora y que mide el avance más que lo logrado? Las concepciones de los docentes sobre esta cuestión, como hemos visto, dependen del contexto en el que se desarrolla su trabajo. De ahí podemos deducir que serán diferentes también sus prácticas evaluativas.

5. DISCUSIÓN

En esta investigación hemos obtenido evidencias empíricas claras de que el contexto socio-económico en el cual se desarrolla la acción docente determina las concepciones evaluativas de los profesores y profesoras. Se han encontrado diferencias en las concepciones sobre los referentes de la evaluación, sobre las pruebas utilizadas y su adaptación, sobre la devolución de información. Todo ello para llegar a configurar una concepción de "evaluación justa" radicalmente diferente en un grupo de docentes y en otro.

Si relacionamos estos resultados con las dos grandes concepciones de evaluación que aborda Remesal (Coll y Remesal, 2009; Remesal, 2011), la más "pedagógica" y la más "social" (y la mixta), se observa una tendencia a tener una visión más social los docentes de contextos favorables y más pedagógica los de contextos desafiantes; eso así dicho con todas las precauciones.

También se han aportado evidencias del gran salto entre los deseos y la realidad, al menos en la evaluación de los estudiantes, tal y como encontró Winterbottom y sus colaboradores (2008). Efectivamente, todos los docentes afirmaron que su práctica evaluativa real dista mucho de lo que ellos consideran que es una evaluación óptima del aprendizaje de los estudiantes.

Sin embargo, lo más sugerente de este estudio es la aparición del concepto de "Evaluación justa" y cómo es visto de forma diferente por los docentes de contextos favorables y desafiantes. El tema de la "justicia", entendida como "*fairness*", es uno de los temas más importantes y que ha atraído a muchos investigadores del ámbito anglosajón (Gipps y Stobart, 2009; Pettifor y Saklofske, 2012), aunque en muchas ocasiones asociado a calidad de los test y su construcción. Menos investigación hay sobre las concepciones de lo que es una evaluación justa por parte de los docentes.

Estos resultados abren una interesante línea de trabajo futura sobre qué consideren los docentes y una evaluación justa y sobre todo sobre los factores que hacen que se construya esa concepción y qué repercusiones tiene para su práctica. La investigación de Scott, Webberb, Lupartc, Aitkend y Scott (2014) es una buena orientación al respecto.

6. CONCLUSIONES

El estudio de las concepciones de los docentes tiene un indudable valor para mejorar las prácticas docentes. Conociendo y comprendiendo cómo son y cómo se construyen esas concepciones es posible modificarlas. Sin embargo esta investigación ha demostrado que las concepciones de los docentes, en muchas ocasiones, tienen mucho más "sentido común" y valor educativo que las volubles obsesiones administrativas.

Este trabajo ha abierto una línea interesante de investigación relacionado con la incidencia del contexto socio-económico en las concepciones de los docentes. Y las conclusiones son claras, dicho contexto determina las concepciones y, con ello, las prácticas y el desarrollo de los estudiantes. De esta forma, frente al supuesto tratamiento indiferenciado por parte de las Administraciones de todos los centros, la investigación muestra que la segregación escolar configura enseñanzas muy diferentes en función del contexto socio-económico. Si tenemos bien asumida la idea de que hay que dar un mayor apoyo al estudiante que más lo necesita, no hay razones para no extrapolar esa idea a los centros.

Esta investigación ha mostrado que las concepciones de los docentes que trabajan en contextos desafiantes tienen elementos de una evaluación más inclusiva, mas adaptada a todos los estudiantes, mientras que los docentes de centros en

entornos más favorables son más exigentes, buscando más la excelencia. Pretender calificar una concepción superior a la otra es tan errado como pensar que esta adaptación al contexto es intrínsecamente positiva. El final, es el contexto el que está determinando el tipo de evaluación, de enseñanza y con ello el futuro de los niños y las niñas.

Parece necesario seguir profundizando en estas ideas realizando un estudio *ex post facto* (siguiendo los estudios de Brown 2003) que permitiera generalizar los resultados. Otra línea de investigación sería el estudio de las concepciones de evaluación de los estudiantes y analizar su relación con los docentes, trazando así nuevas preguntas de investigación para la mejora del aprendizaje de los estudiantes y la enseñanza de los docentes.

La forma en que los docentes conciben su evaluación y cómo ésta se traduce en su práctica diaria, es un elemento clave para la mejora educativa. Ser conscientes de cómo se evalúa y qué pasos faltan para llegar a una mejor evaluación es uno de los retos del sistema educativo, así como implicar más a los estudiantes en su proceso de evaluación y de aprendizaje. La importancia del contexto en la construcción de esas concepciones nos abre la puerta a la necesidad de dar un tratamiento diferencial a las escuelas teniendo en consideración su contexto. El debate sobre qué es una escuela justa, cómo se construye esa idea y qué implicaciones tiene, es un tema que deberíamos abordar si queremos una evaluación que genere una educación que haga una sociedad más justa y equitativa (Murillo e Hidalgo, 2014, 2015).

REFERENCIAS BIBLIOGRÁFICAS

- Atjonen, P. (2014). Teachers' views of their assessment practice. *The Curriculum Journal*, 25(2), 238-259.
- Azis, A. (2012). Teachers' conceptions and use of assessment in student learning. *Indonesian Journal of Applied Linguistics (IJAL)*, 2(1), 40-52.
- Black, P., Harrison, C., Lee, C., Marshall, B., y William, D. (2002). *Working inside the black box: Assessment for learning in the classroom*. Londres: Department of Education and Professional Studies, Kings College.
- Bouden, J.A., y Walsh, E. (2000). *Phenomenography*. Victoria: Royal Melbourne Institute of Technology University Press.
- Brown, G.T.L. (2003). *Teachers' conceptions of assessment*. Auckland: University of Auckland.
- Brown, G.T.L. (2004). Teachers' conceptions of assessment: implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301-318.
- Brown, G.T.L., y Hirschfeld, G.H. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3-17.
- Brown, G.T.L., y Remesal, A. (2012). Prospective Teachers' Conceptions of Assessment: A Cross-Cultural Comparison. *The Spanish journal of psychology*, 15(1), 75-89.
- Brown, G.T.L., Lake, R. y Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and*

Teacher Education, 27(1), 210-220.

Chen, J., Brown, G.T.L., Hattie, J.A.C., y Millward, P. (2012). Teachers' conceptions of excellent teaching and its relationships to self-reported teaching practices. *Teaching and Teacher Education*, 28(7), 936-947.

Coll, C., y Remesal, A. (2009). Concepciones del profesorado de matemáticas acerca de las funciones de la evaluación del aprendizaje en la educación obligatoria. *Infancia y aprendizaje*, 32(3), 391-404.

Fives, H., y Gill, M.G. (2014). *International Handbook of Research on Teachers' Beliefs*. Londres: Routledge.

Gipps, C., y Stobart, G. (2009). Fairness in assessment. En C. Wyatt-Smith y J. Cumming (Eds.), *Educational assessment in 21st century: Connecting theory and practice* (pp. 105-118). Dordrecht: Springer.

Griffiths, T., Gore, J., y Ladwig, J. (2006). Teachers' fundamental beliefs, commitment to reform and the quality of pedagogy. *Conferencia anual de la Asociación australiana de investigación en educación*. Universidad Adelaide, Australia.

Harris, L., y Brown, G.T.L. (2009). The complexity of teachers' conceptions of assessment: tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practices*, 16, 363-381.

Kvale, S. (1996). *Interviews: an introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage Publications.

Lupton, R. (2005). Social justice and school improvement: improving the quality of schooling in the poorest neighbourhoods. *British Educational Research Journal*, 31(5), 589-604.

- Marshall, B., y Drummond, M.J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education* 21(2), 133–49.
- Marton, F. (1986). Phenomenography: A research approach to investigating different understandings of reality. *Journal of Thought*, 21(3), 28-49.
- Marton, F., y Booth, S. (1997). *Learning and Awareness*. NuevaYork: Lawrence Erlbaum.
- Murillo, F.J., e Hidalgo, N. (2014). Evaluación justa: entre la igualdad y la equidad. *Revista Iberoamericana de Evaluación Educativa*, 7(3), 5-6.
- Murillo, F.J., e Hidalgo, N. (2015). Dime cómo evalúas y te diré qué sociedad construyes. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 5-9.
- Murillo, F.J., Martínez-Garrido, C., e Hidalgo, N. (2014). Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España. *Estudios sobre Educación*, 27, 91-113.
- Orgill, M.K. (2002). *Phenomenography*. Recuperado de <http://chemed.chem.purdue.edu/chemed/bodnergroup/frmaeworks/phenomenography.htm>
- Pajares, M.F. (1992). Teachers' Beliefs and Educational Research: Cleaning Up a Messy Construct. *Review of Educational Research*, 62(3), 307-332.
- Pettifor, J.L., y Saklofske, D. H. (2012). Fair and ethical student assessment practices. En C.F. Webber y J. Lupart (Eds.), *Leading student assessment* (pp. 87–106). Dordrecht: Springer.
- Pope, N., Green, S. K., Johnson, R. L., y Mitchell, M. (2009). Examining teacher ethical

- dilemmas in classroom assessment. *Teaching and Teacher Education*, 25, 470-483.
- Pratt, D.D. (1992). Conceptions of teaching. *Adult Education Quarterly*, 42(4), 203-220.
- Prieto, M. (2008). Creencias de los profesores sobre Evaluación y Efectos Incidentales. *Revista de Pedagogía*, 29(84), 123-144.
- Prieto, M., y Contreras, G. (2008). Las concepciones que orientan las prácticas evaluativas de los profesores: un problema a develar. *Estudios Pedagógicos*, 34(2), 245-262.
- Reed, B. (2006). Phenomenography as a way to research the understanding by students of technical concepts. *Technological Innovation and Sustainability*, 1, 1-11.
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472-482.
- Richardson, J.T.E. (1999). The Concepts and Methods of Phenomenographic Research. *Review of Educational Research*, 69(1), 53-82.
- Rubie-Davies, C.M. Flint, A., y McDonald, L.G. (2012). Teacher beliefs, teacher characteristics, and school contextual factors: What are the relationships? *British Journal of Educational Psychology*, 82(2), 270-288.
- Scott, S., Webberb, C., Lupart, J., Aitkend, N., y Scott, D. (2014). Fair and equitable assessment practices for all students. *Assessment in Education: Principles, Policy & Practice*, 21(1), 52-70.
- Thompson, A.G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. En D.A. Grouws (Ed.), *Handbook of research on mathematics*

teaching and learning (pp.127-146). Nueva York: Macmillan.

Trigwell, K. (2000). A phenomenographic interview on phenomenography. En J. A.

Bowden y E. Walsh (Eds.), *Phenomenography* (pp. 62-82). Victoria: RMIT.

Van den Berg, R. (2002). Teachers' meanings regarding educational practice. *Review of Educational Research*, 72(4), 577-625.

Winterbottom, M., Brindley, S., Taber, K.S., Fisher, L.G., Finney, J., y Riga, F. (2008).

Conceptions of assessment: trainee teachers' practice and values. *The Curriculum Journal*, 19(3), 193-213.

Vol. 25, Num. 1 (ENERO-JUNIO) 2014

ISSN 1130-2496 ISSNe 1988-2793

A1.4. Publicación 4.

**Las concepciones sobre una evaluación justa de los
estudiantes. Un estudio fenomenográfico desde la
perspectiva de los docentes**

Revista Complutense
de **Educación**

PUBLICACIONES UNIVERSIDAD
COMPLUTENSE DE MADRID



Revista Complutense de Educación

Facultad de Educación - Centro de Formación del Profesorado
Universidad Complutense de Madrid
e-mail: rce@edu.ucm.es

DOÑA COVADONGA RUIZ DE MIGUEL, DIRECTORA DE LA REVISTA
COMPLUTENSE DE EDUCACIÓN, CON ISSN 1130-2496.

CERTIFICA

Que el artículo titulado " LAS CONCEPCIONES SOBRE UNA EVALUACIÓN JUSTA DE LOS ESTUDIANTES. UN ESTUDIO FENOMENOGRAFICO DESDE LA PERSPECTIVA DE LOS DOCENTES" elaborado por F. JAVIER MURILLO Y NINA HIDALGO FARRAN, ha sido evaluado favorablemente por el Consejo Evaluador de la Revista Complutense de Educación y se publicará en uno de los próximos números de la Revista.

Y, para que conste a los efectos oportunos, firma el presente certificado en Madrid, a 15 de febrero de 2017.



Covadonga Ruiz de Miguel
Editora
Revista Complutense de Educación

Las concepciones sobre una evaluación justa de los estudiantes. Un estudio fenomenográfico desde la perspectiva de los docentes

Resumen. La presente investigación se centra en comprender las concepciones de los docentes acerca de qué es una evaluación justa. Las concepciones de los profesores son las creencias implícitas que tienen sobre el proceso de enseñanza-aprendizaje y tienen una fuerte incidencia sobre sus prácticas evaluativas y sobre el desarrollo de los estudiantes. Para lograr nuestro objetivo hemos desarrollado un estudio fenomenográfico. Los participantes son 31 docentes de Educación Primaria y de Educación Secundaria de la Comunidad de Madrid. Los resultados muestran que las concepciones de evaluación justa de los docentes están divididas, estando algunas más cercanas al principio de igualdad y otras al de equidad. La conclusión final del estudio es la influencia del contexto del centro educativo en la concepción de evaluación justa del docente y la relevancia de conocer las concepciones de los docentes para poder implementar prácticas evaluativas que logren una educación y una sociedad más justa.

Palabras clave: evaluación de estudiantes, concepciones docentes de evaluación, investigación fenomenográfica, evaluación justa, investigación cualitativa.

[en] Fair assessment conceptions of students. A phenomenographic study from teachers' perspective

Abstract. The present research focuses on understanding teachers' conceptions about what is a fair assessment. Teachers' conceptions are the implicit beliefs that they have about the teaching-learning process and have a strong impact on their assessment practices and student development. To achieve our purpose, we have developed a phenomenographic study. The participants are 31 teachers of Primary and Secondary Education of the Community of Madrid. The results show that teachers' conceptions of fair assessment are divided, some of them being closer to the principle of equality and others to equity. The final conclusion of the study is the influence of the context of the school on fair assessment' conceptions of the teachers and the relevance of knowing that if we understand teachers' conceptions we are going to be able to implement assessment practices that achieve an education and a more just society.

Keywords: student assessment, conceptions of assessment, phenomenographic research, fair assessment, qualitative research.

Sumario. 1. Introducción. 2. Marco Teórico. 3. Método. 3.1. Enfoque metodológico. 3.2. 3.2. Participantes e instrumento de recogida de datos. 3.3. Procedimiento y análisis de datos. 4. Resultados. 4.1. Concepción igualitaria. 4.2. Concepción Equitativa. 5. Discusión y conclusiones. 6. Referencias.

1. Introducción

En una investigación ya clásica, Sambell, McDowell y Brown (1997) afirmaron que la evaluación del aprendizaje de los estudiantes es uno de los procesos educativos con una percepción más negativa por los estudiantes y docentes. Así, encontraron que los agentes del proceso educativo conciben la evaluación como un estorbo, un mal necesario o una injusticia; también que los instrumentos y estrategias de evaluación relacionadas con la certificación así como los exámenes son concebidos como prácticas ajenas al aprendizaje real de los estudiantes.

Con esos antecedentes no es extraño que en las últimas décadas se hayan multiplicado las investigaciones que buscan conocer y explicar las concepciones de profesores y profesoras sobre la evaluación (Azis, 2015; Harris y Brown, 2009; entre otros). La idea es sencilla, en la medida que las concepciones marcan la práctica, si es posible conocer las concepciones de los docentes, se tendrá más información para modificar sus acciones y, con ello, optimizar el aprendizaje de los estudiantes (Griffiths, Gore y Ladwig, 2006; Autores, 2014).

Si hablamos de una evaluación justa, el interés de los estudiosos es más que evidente. Solo podemos avanzar en el desarrollo de una evaluación más justa si, conociendo las concepciones de los docentes, se trabaja sobre ellos. De ahí viene el interés despertado en estos últimos años (Pettifor y Saklofske, 2012; Sambell, McDowell y Brown, 1997; Suskie, 2002). Sin embargo, las investigaciones empíricas sobre las concepciones de los docentes hacia una evaluación justa son muy escasas, lo que justifica el interés de esta investigación.

Además, hay que considerar los acelerados cambios en las prácticas evaluativas que está viviendo el sistema educativo español, contexto donde se sitúa la presente investigación. En estos momentos se está comenzando la implantación generalizada de pruebas estandarizadas externas con carácter informativo y certificativo tanto en Educación Primaria como en Secundaria que son tachadas de injustas por muchos profesores. Este hecho ha abierto el debate acerca de los justo o injusto de las mismas y acerca de a quién beneficia este enfoque.

2. Marco teórico

Entendemos por concepciones un sistema organizado de creencias originadas en las experiencias del individuo y desarrolladas a través de las interacciones en las que participa (Coll y Remesal, 2009; Remesal, 2011). Las concepciones tienen un fuerte carácter social, ya que se fundamentan y construyen en espacios compartidos, viéndose influidas por las interacciones de la persona con el mundo (Van den Berg, 2002). Aplicado al ámbito educativo, las concepciones de los docentes se entienden como una estructura de creencias a través de la cual el profesor interactúa diariamente en su aula (Prieto y Contreras, 2008). Dichas concepciones, también se ven influidas por el contexto en el que se encuentra el docente, definiendo su práctica profesional. Las concepciones del docente son, sin duda, un elemento determinante en el proceso de enseñanza-aprendizaje y con ello en la evaluación (Griffiths, Gore y Ladwig, 2006).

Uno de los investigadores que es un referente en la línea de investigación de las concepciones sobre evaluación de los docentes es Gavin T. L. Brown (Brown, 2003, 2006,

2008; Brown y Gao, 2015; Gebril y Brown, 2014). En los últimos años, él y sus colegas (Lois Ruth Harris, Gerrit Hirschfeld, Kerry Kennedy, Elizabeth Peterson y Earl Irving, entre otros) desarrollaron una ambiciosa investigación para determinar las concepciones sobre evaluación de profesores y estudiantes y sus implicaciones. Los datos fueron recolectados a través del *Teachers' Conceptions of Assessment* (COA-III)) y analizados a través de modelos de ecuaciones estructurales, mostrando un ajuste cercano de los datos a un modelo multidimensional y jerárquico.

En este modelo, Brown (2003, 2006) muestra que existen cuatro tipos de concepciones de los maestros sobre la evaluación: a) la evaluación fomenta la rendición de cuentas de la escuela, ya que proporciona información sobre el funcionamiento de los colegios; b) la evaluación hace que los estudiantes rindan cuentas, midiendo su desempeño académico c) la evaluación mejora la educación, ya que proporciona información clara y valiosa sobre el desarrollo y progreso de los estudiantes y su aprendizaje, y d) la evaluación es irrelevante, ya que es injusta para los estudiantes y aporta información incorrecta y poco útil para la enseñanza. Los análisis realizados por los investigadores revelan que las concepciones sobre evaluación de los docentes se mantienen constantes y estables y claramente relacionadas con la frecuencia que utilizan las distintas tareas de evaluación en el aula (Harris y Brown, 2008).

Profundizando en la relación entre las diferentes concepciones de evaluación de los profesores, los resultados de la investigación determinan que los maestros que están de acuerdo con la concepción de evaluación como una herramienta de mejora también están moderadamente de acuerdo en que la evaluación debe ser un instrumento para la rendición de cuentas de la escuela. Los profesores que se mostraron menos cercanos a una concepción de la evaluación como irrelevante, mostraban una alta correlación negativa con la concepción de evaluación como una herramienta de mejora del aprendizaje ($r_{xy} = -0,69$). Por último, las concepciones docentes de evaluación referentes a rendición de cuentas tanto de la escuela como de los estudiantes correlacionan positivamente ($r_{xy} = 0,58$) (Brown, 2008).

En el ámbito nacional, se han empezado a desarrollar algunas investigaciones que han evidenciado la importancia de profundizar en las concepciones sobre la evaluación de los estudiantes para mejorar, de un lado el diseño y la tarea evaluativa de los docentes y por otro cómo ésta puede favorecer el aprendizaje del alumnado (Brown y Remesal, 2012; Coll y Remesal, 2009; Prieto y Contreras, 2008; Remesal, 2011).

Dentro de la línea de investigación sobre las concepciones docentes de la evaluación, cobra especial relevancia el estudio de las concepciones sobre qué es una evaluación justa (Cole y Zieky, 2001; Gipps y Stobart, 2009, Pettifor y Saklofske, 2012). Definir qué es una evaluación justa no es una tarea fácil, ya que tiene implicaciones éticas que influyen enormemente en el aprendizaje de los estudiantes. Uno de los elementos que más dificultan la conceptualización de una evaluación justa es el binomio igualdad/equidad. Así, la evaluación justa se concibe más próxima a la igualdad o más cercana a la equidad.

Lam (1995) indica que los docentes que conciben una evaluación justa desde una perspectiva igualitaria se centran en garantizar la igualdad de condiciones en la realización de la prueba (materiales, tiempo, espacio, recursos, etc.). La evaluación igualitaria requiere la misma administración, contenidos, puntuación e interpretación de los resultados, persiguiendo, por lo tanto, la objetividad.

Frente a esta postura, autores como Suskie (2002) defienden que una evaluación justa no es aquella que es igual para todos los estudiantes, ya que no todos han tenido necesariamente las mismas oportunidades de acceso a la educación; por lo que el principio de igualdad no es suficiente para garantizar la justicia. Suskie (2002) considera esencial la promoción de la equidad en la evaluación para garantizar la justicia, adaptando la evaluación a las necesidades y características de los estudiantes. Esta equidad se traduce en llevar a cabo múltiples evaluaciones y con diferentes instrumentos para asegurar que la diversidad de los estudiantes se atiende correctamente a través de la evaluación.

Picón-Jácome (2013), por su parte, defiende que una evaluación justa solo puede entenderse desde una perspectiva equitativa a través del desarrollo de una evaluación de los estudiantes basada en principios democráticos, promoviendo la participación de los alumnos y alumnas en la evaluación y siendo una herramienta útil para ayudar a los estudiantes en la mejora de su aprendizaje.

Reflexionando acerca del binomio igualdad-equidad, Lam (1995) considera que una evaluación es injusta si (1) a los alumnos no se les brinda igualdad de oportunidades para demostrar lo que saben, por lo que las evaluaciones son parciales, (2) estas evaluaciones parciales se utilizan para juzgar las capacidades y necesidades de los alumnos, y (3) si los docentes tienen una visión distorsionada del estudiante y esta visión se traduce en la toma de decisiones educativas que finalmente limita las oportunidades educativas del alumno.

Complementariamente, Pettifor y Saklofske (2012) afirman que existen diversas cuestiones pedagógicas que transforman a la evaluación en una práctica más justa. Una de estas cuestiones es que los estudiantes conozcan los criterios de evaluación con los que van a ser evaluados, o incluso mejor, que dichos criterios sean co-definidos entre los docentes y los estudiantes (Andrade y Du, 2007; Harlen, 2005; Pepper y Pathak, 2008). Otro elemento importante para conseguir una evaluación más justa es no discriminar a los estudiantes por sus características sociales, culturales o económicas (que suelen tener impacto en la evaluación) (Stobart, 2005). Los autores también remarcan que los docentes no solamente evalúan las pruebas escritas sino también la personalidad, progreso, esfuerzo y participación de los estudiantes, ya que solo añadiendo estos aspectos más cualitativos a la evaluación se logrará una práctica evaluativa justa. Asimismo, es esencial que la evaluación sea individualizada a las necesidades de cada estudiante.

Otro estudio que arroja luz a esta cuestión es el estudio llevado a cabo por Tierney, Simon y Charland (2011) en Canadá. El propósito del estudio se centró en determinar cómo 77 docentes evaluaron a sus estudiantes en dos sistemas educativos canadienses diferentes con distintas políticas y estándares de evaluación. Analizando los resultados, los autores concluyen que los docentes evalúan teniendo en cuenta lo que creen que es justo para sus estudiantes, por el bien individual de cada estudiante y común del aula y de la escuela.

Como hemos señalado, uno de los elementos más llamativos sobre las concepciones de una evaluación justa, es que existen pocas investigaciones empíricas que arrojen datos sobre el tema. El trabajo de Scott, Webber, Lupart, Aitken y Scott (2014) es una buena orientación al respecto. Estos autores realizaron un estudio de enfoque mixto (cuantitativo-cualitativo) con el objetivo de profundizar en las concepciones de evaluación justa de los docentes. Los participantes fueron 3.312 docentes, estudiantes, familias y equipo directivo. Los resultados muestran que las prácticas evaluativas justas son más próximas al concepto de equidad y

ofrecen cinco ideas clave sobre cómo debe ser la práctica evaluativa justa: a) Los educadores deben esforzarse por comprender el impacto que suponen sus prácticas de evaluación a los estudiantes y a sus familias; b) la evaluación debe ser diferenciada a cada estudiante, para atender a la capacidad, nivel socioeconómico, cultura y lengua de los estudiantes; c) todos los miembros de la comunidad educativa es necesario que denuncien prácticas indefendibles e ilógicas de evaluación; d) la frecuencia, la intensidad y la intrusión de las evaluaciones no debe abrumar ni a los estudiantes ni a sus familias; y e) la evaluación no debe ser utilizada para contrarrestar el comportamiento inapropiado del estudiante o como recompensa deseada a un determinado comportamiento.

Con todo ello, la presente investigación tiene como objetivo conocer las concepciones de Evaluación Justa de docentes de Educación Primaria y Secundaria.

3. Método

3.1. Enfoque metodológico

El enfoque metodológico más adecuado para profundizar en las concepciones de las personas, en este caso de los docentes, es la Fenomenografía (Bowden y Walsh, 2000; Marton, 1986). La misma se define por tener el propósito de “identificar de forma cualitativa las diferentes formas en que las personas experimentan, conceptualizan, perciben y comprenden varios aspectos del fenómeno y de su mundo próximo” (Marton, 1986, p. 31), y cuyo objetivo es “...descubrir y sistematizar formas de pensamiento que sintetizan el modo en que las personas interpretan aspectos de la realidad” (Marton, 1986, p. 180).

En la investigación fenomenográfica se hace un esfuerzo por descubrir todas las comprensiones que la gente tiene sobre los fenómenos específicos y ordenarlos en categorías conceptuales, partiendo de la idea de que las concepciones son la forma central del conocimiento humano, por lo que la concepción es la unidad básica de la descripción en la investigación fenomenográfica.

3.2. Participantes e instrumento de recogida de datos

Un punto clave en la investigación fenomenográfica es la selección de los participantes. No hay que perder de vista que la finalidad de los estudios fenomenográficos es contar con la mayor variación posibles de una misma concepción, por lo que los sujetos tienen que ofrecer distintas formas de entender y percibir una realidad. Los participantes en esta investigación son 31 profesores de Educación Primaria y de Educación Secundaria que trabajan en centros públicos, privados concertados y privados no concertados de la Comunidad de Madrid. La selección de los participantes se hizo mediante un muestreo no probabilístico por cuotas con tres criterios de selección (tabla I):

- *Contexto socio-económico del centro educativo*: hemos dividido los centros entre aquellos situados en contextos favorables y los situados en contextos desafiantes. Los indicadores que nos han permitido seleccionar estos centros han sido: el nivel socio cultural del barrio donde se encuentra el centro, el nivel socio-económico de las familias de los estudiantes que asisten a ese centro y las características sociales, culturales y lingüísticas de dichos estudiantes.

- *Nivel educativo del centro educativo*: este indicador se refiere al nivel educativo en el se encuentra, diferenciado entre los docentes de Educación Primaria y los de Educación Secundaria.
- *Titularidad del centro*: también se ha tenido en cuenta la titularidad del centro, contando con docentes que trabajan en centros educativos públicos, concertados y privados.

Contexto	Educación Primaria		Educación Secundaria			Total
	C. públicos	C. concertados	C. públicos	C. concertados	C. privados no concertados	
Desafiante	3	3	4	4		14
Favorable	4	5	4	2	2	17
Total	7	8	8	6	2	31

Tabla I. Distribución de los participantes en función del contexto, del nivel educativo y de la titularidad del centro

El proceso de selección de los participantes fue el siguiente. En primer lugar se seleccionaron zonas con un alto estatus socio-económico -contexto favorable- y de bajo estatus -contexto desafiante- en la Comunidad de Madrid. En cada una de las zonas se seleccionaron centros de Educación Primaria y de Educación Secundaria públicos y privados. Se contactó personalmente con el equipo directivo de cada escuela elegida, y con su ayuda se seleccionaron varios profesores y profesoras para participar en el estudio. Así, se realizaron entre 2 y 3 entrevistas en cada centro. De ellos, 19 son mujeres y 11 hombres (10/5 en Educación Primaria y 9/6 en Educación Secundaria), con un promedio de 13 años de experiencia docente.

El principal instrumento de recogida de información ha sido la entrevista fenomenográfica (Trigwell, 2000) basada en el auto-relato (Marton y Booth, 1997). Las entrevistas partieron de una pregunta generadora: *¿qué es para ti una evaluación justa?*

La entrevista fenomenográfica es semi-estructurada, siendo necesario que el guión apenas consta de algunas preguntas clave predeterminadas (Bowden, 1996; Marton, 1986). Es muy relevante que el investigador tenga en cuenta el foco del estudio, ya que éste es el que enfoca y guía la entrevista en todo momento. En este caso, el investigador ha profundizado a través de las entrevistas de las concepciones de los docentes acerca de que es una evaluación justa y sus implicaciones. Asimismo, el investigador no debe posicionarse frente a un determinado tipo de concepción, sino ser neutral y mostrar interés compartido con todos los participantes.

Cada una de las entrevistas realizadas a los docentes ha tenido una duración de entre veinte minutos a media hora. Una vez realizadas todas las entrevistas se transcribieron por completo y se etiquetó cada una de ellas en función de las características de la escuela en la que trabaja cada docente de acuerdo al siguiente código: [D1, SEC, PR, FAV:3]. En la etiqueta, podemos diferenciar las siguientes informaciones: a) D1: número de docente, b) nivel educativo (PRI: educación Primaria y SEC: educación Secundaria, c) titularidad del centro (PU: público y PR: privado), d) contexto donde se ubica el centro educativo (FAV:

favorable y DES: desafiante) y, por último, e) número de la página de la entrevista transcrita donde se encuentra la cita.

Para evitar los sesgos tanto en la recogida de datos como en el análisis de los mismos los investigadores tuvieron ciertos aspectos éticos que guiaron su trabajo. Así, los participantes fueron informados de los objetivos de la investigación, se solicitó su consentimiento para ser grabados; se mantuvo escrupulosamente su anonimato y fuimos extremadamente cuidadosos de no alterar ni sus palabras ni sacarlas del contexto.

El trabajo de campo fue desarrollado entre los meses de mayo a diciembre de 2014. El proceso fue el siguiente: una vez contactadas las escuelas e informadas del propósito de la investigación, acudimos a cada uno de ellas y tuvimos una breve reunión con el equipo directivo. A continuación, el equipo directivo facilitó un encuentro con los posibles docentes a entrevistar en cada centro, donde se profundizó en la entrevista fenomenográfica a realizar y de la disponibilidad de cada docente. En esa reunión, los docentes firmaron un consentimiento de acuerdo al conocimiento del propósito de la investigación, el anonimato de las entrevistas y la fidelidad de los datos.

3.3. Procedimiento y análisis de datos

De acuerdo con el método fenomenográfico se ha utilizado un criterio de proximidad al discurso de los participantes. Una vez realizadas las entrevistas se analizó, con el software Atlas.ti 7, la presencia de diferentes concepciones pertenecientes a determinadas categorías y sub-categorías en el discurso docente.

Para el análisis de los datos llevamos a cabo un procedimiento similar al propuesto por Marton (1986). En primer lugar, se identificaron las concepciones preconcebidas antes de analizar los datos. En segundo lugar, los investigadores nos centramos en las palabras de los docentes, dejando de lado las perspectivas teóricas sobre una “evaluación justa”. Las categorías creadas surgieron de las propias palabras de los docentes, favoreciendo que el análisis sea lo más fiel posible a las concepciones de los profesores y profesoras entrevistados. En tercer lugar, comparamos las concepciones de un docente con el resto, viendo regularidades y creando pre-categorías cada vez más amplias, analizando asimismo la relación entre las distintas pre-categorías (Marton, 1986).

De acuerdo con Sjostrom y Dahlgren (2002), basamos el análisis de las pre-categorías de acuerdo a tres principios:

- *Frecuencia*, relativo a cuantas veces los docentes mencionan esa idea o concepción relacionada con una evaluación justa.
- *Posición*, donde se ubica la concepción respecto al resto de concepciones, identificando aquellas más relevantes.
- *Pregnancy*, analizando aquellos aspectos en los que los docentes entrevistados ponen más énfasis o consideran más importante para desarrollar una evaluación justa.

Una vez seleccionadas las pre-categorías relacionadas con una evaluación justa, se seleccionaron aquellas citas más representativas de cada una de las categorías, siendo especialmente cuidadosos con respetar el sentido de las ideas expresadas por los docentes

entrevistados. Por último, las pre-categorías se analizaron en distintas redes de significado¹. Estas redes de significado fueron analizadas de forma individual y después de forma cruzada, estableciendo las categorías finales y representando las distintas formas de concebir la evaluación justa.

Siguiendo las características de la investigación fenomenográfica, esta explicación detallada del proceso de análisis, así como la explicitación de los referentes teóricos, garantizan la calidad, adecuación y transferibilidad de los resultados encontrados (Åkerlind, 2002; Harris, 2008).

4. Resultados

Las entrevistas fenomenográficas realizadas arrojan una visión llena de matices que nos ayuda a profundizar en las concepciones de los profesores sobre qué es una evaluación justa. Los hallazgos muestran que, esencialmente, existen dos concepciones sobre qué es una evaluación justa: una que vincula la justicia con la igualdad; y otra que la relaciona con la equidad. En muchos casos, los docentes se ubican con claridad en una u otra concepción; pero también se encuentran casos con concepciones contradictorias o cambiantes.

4.1. Concepción igualitaria

Un grupo de docentes tienen una concepción igualitaria de lo que es evaluación justa. Estos profesionales destacan cuatro elementos fundamentales: la transparencia en la información, la objetividad de las pruebas, el carácter eminentemente cuantitativo de la evaluación, y que contribuyan a capacitar a los estudiantes a superar exitosamente las evaluaciones externas a las que serán sometidos en su vida académica.

En primer lugar, los docentes consideran que para que una evaluación sea justa es necesario transparencia. Es decir, que todos los estudiantes conozcan el momento, los contenidos, los instrumentos y los criterios que el profesorado va a utilizar para evaluar a los estudiantes. En palabras de dos docentes:

Es justo que se sepa de qué te van a evaluar, qué cae en el examen, cómo va a ser la prueba, qué es lo que se va... lo que se espera que hagas para sacar buena nota...
[D8,SEC,CC,FAV: 4]

Yo creo que un componente importante de la justicia es la honestidad, entonces yo creo que tienen que haber una claridad en la evaluación y saber a qué atenerse; tener unos criterios establecidos y en función de esos criterios la persona que es evaluada sabe de qué va a ser evaluada y sabe para qué es la evaluación. [D7,SEC,CP,FAV: 3]

Los docentes consideran necesario que haya una absoluta claridad en las calificaciones obtenidas, lo que refuerza la idea de que los criterios sean explícitos e iguales para todos. Siguiendo esta idea, un docente indica que:

“Siempre tiene que haber unos criterios de calificación expresos y los conocimientos que tienen que aprender. Tienen que estar enterados de donde están y de lo que se les está pidiendo”.
[D21,SEC,CC,DES: 4]

¹ Del inglés *pool of meaning*.

En segundo lugar, para los docentes una evaluación justa es una evaluación objetiva. De esta manera, la forma más justa de evaluar a los estudiantes es aquella que no ofrece margen de error, siendo absolutamente cerrada y que no admita discusión o debate. Un docente nos señala lo siguiente:

Yo lo que busco es la objetividad [...]. Corrijo lo más objetivamente posible. Mira, la de verdadero y falso es muy fácil, 25 por 0,4. [...] Me gusta de tal manera que esté estructurado y que ellos sepan. [D8,SEC,CC,FAV:3]

Así, la objetividad y la transparencia son dos criterios que debe tener una evaluación justa. Tal como indican los docentes:

Una evaluación en la que se tengan en cuenta todo: el entorno del alumno, todas sus características, su evolución, el proceso que ha tenido... [D5,PRI,CPU,FAV:4]

Estas ideas están íntimamente ligadas a la importancia de una evaluación cuantitativa como forma de visibilizar la objetividad. La concepción que tienen estos docentes es que este tipo de pruebas son más objetivas y posibilitan tener información más clara a los estudiantes para mejorar su aprendizaje. Así, los docentes indican que el uso de pruebas objetivas como el examen ofrecen una visión adecuada del aprendizaje de los estudiantes. En esta línea, un profesor indica que:

Hacerles cuatro exámenes seguidos, lo siento pero es lo que toca. ¿Por qué? Pues porque si tú les pones tres ejercicios en un examen de un temario que tiene 20 o 25 aspectos no tengo argumentos como para poder evaluar a los alumnos. Objetivamente, tienes más probabilidades de que demuestren lo que saben. Así cuando pongo la nota, estoy seguro al 90% de que refleja lo que sabe. [D13,SEC,CPU,FAV:3]

Un último elemento que destacan los docentes al preguntarles que es una evaluación justa, es que ayude a los estudiantes a prepararles para superar las pruebas objetivas (CDI de Madrid, Pruebas de Acceso a la Universidad, etc.) a las que tendrán que hacer frente a lo largo de su vida académica. En este caso, la justicia es concebida en términos de preparar a los alumnos a hacer las pruebas externas de la mejor manera posible ya que eso influirá enormemente en el camino de vida elegido por el estudiante. En palabras de una maestra de Educación Primaria:

Además están ahí pendientes una serie de evaluaciones que van a ser determinantes para los chavales, para elegir sus caminos de formación y tal, y eso va a influir en cómo evaluemos también en el centro porque al final se nos va a pedir un poco desde los padres y desde la sociedad, que esos niños aprueben esas pruebas para poder seguir formándose y vivir. [D27,PRI,CC,FAV:4]

Resulta interesante encontrar que en el grupo de docentes encontramos profesores que imparten clase tanto de Educación Primaria, como de Educación Secundaria y que se encuentran en centros titularidad pública y privada (concertada o no concertada); sin embargo, en su práctica totalidad son profesores caracterizados por trabajar en centros educativos situados en contextos socio-económicamente favorables.

4.2. Concepción equitativa

Por otro lado, existe otro grupo de docentes que consideran que una evaluación justa es

aquella que se basa en la equidad. Estos profesionales destacan que es necesario que dicha evaluación se adapte a las necesidades de los estudiantes y tenga en cuenta su contexto, que mida realmente el avance o progreso en el aprendizaje y que sea cualitativa y procesual.

Los profesores definen la evaluación justa como aquella que se centra en cada uno de los estudiantes. Según los docentes, es necesario que la práctica evaluativa se adapte a las características y necesidades de los estudiantes, alejándose por lo tanto de evaluaciones homogéneas. Ilustrando esta idea:

La evaluación tiene que ser mucho más individualizada, tú en concreto, David, María José o Pepito tu evaluación es ésta, por tus características concretas, por tu contexto, porque tú has mostrado más interés, porque tú hayas sacado un 5 en un examen no te preocupes, porque tú no eres María tienes un 5 en inglés, no? Hay muchas otras cosas que hay que tener en cuenta y en fin, hacerles ver que les estamos conformando como personas. [D23,PRI,CPU,DES:4]

Los docentes indican que además de ser adaptada a cada estudiante es necesario que la evaluación tenga en cuenta el contexto del estudiante así como el entorno en el que crece y se desarrolla, ya que éste tiene muchas implicaciones en su desarrollo:

Pues una evaluación en la que se tengan en cuenta todo el entorno del alumno, todas sus características, su evolución, el proceso que ha tenido... [D4,PRI,CC,DES:3]

En segundo lugar, los docentes defienden que una evaluación justa es necesario que sirva para medir el esfuerzo de los estudiantes. Este esfuerzo requiere tener presente el punto de partida de los alumnos y realizar continuas evaluaciones para comprobar su avance en el aprendizaje.

Un profesor indica que:

Una evaluación justa sería aquella evaluación que reflejase el esfuerzo del niño durante esos 15, 20 días o un año... si tú quieres evaluar contenidos concretos o una asignatura, que seas capaz de valorar el esfuerzo que el chico ha realizado. Si no ha habido ningún tipo de esfuerzo porque el chico... porque los contenidos estaban poco adecuados y no ha aprendido nada, esa evaluación está mal hecha, porque no le estás aportando nada al chico. [D2,PRI,CPU,DES:3]

Muy vinculado a este esfuerzo, los docentes también consideran que una evaluación es justa cuando ayuda a que los estudiantes aprendan para su vida, es decir, que dicho aprendizaje sea útil y transferible. Mostrando esta idea, una maestra de Primaria indica que:

Tenga en cuenta el esfuerzo, la motivación y que hayan sacado un poco de partido. No tanto que hayan aprendido los contenidos que tengan que aprender, y eso no sirve. Yo veo si un grupo o un alumno tiene resultados positivos en cuanto a si ha crecido en responsabilidad, ha crecido en un cierto gusto por el aprendizaje, o si les he despertado su curiosidad por aprender. [D22,PRI,CPU,DES:1]

Para otro docente, una evaluación justa valora positivamente (premia) a aquellos estudiantes que han avanzado más en su aprendizaje, mencionando los retos como forma de aprendizaje y evaluación que favorece la justicia en el aula:

Yo evalúo mucho más a una persona que le costó desde el primero momento adquirir ese conocimiento y él se lo ha luchado, se lo ha peleado y ha conseguido adquirir ese conocimiento... o sea para mí no puede tener la misma nota el niño que me ha sacado las

mismas cinco operaciones que desde el primer momento las coge fenomenal, que al niño que le ha costado un mes llegar a ellas. Porque a lo mejor al que le ha salido desde el primer momento fenomenal, podía haber llegado un escalón más arriba... [D3,PRI,CC,DES:3]

Hay que optar a que el premio no sea la nota, porque entonces empiezan las competiciones, lo mejor es que optes por otro tipo de premios, por ejemplo, al final un alumno que lo ha hecho todo bien lo que se merece es un reconocimiento, es aumentar su responsabilidad en el aula. [D28,PRI,CPU,DES:2]

En tercer lugar los docentes consideran que para que una evaluación sea verdaderamente justa debe alejarse de la objetividad, y contar con instrumentos y procesos más cualitativos. Así, reflexionan que es fundamental complementar los instrumentos más cuantitativos (como puede ser un examen) con información de corte más cualitativo, a través de la observación, diarios de campo del profesor o portafolios. Los docentes entrevistados reflexionan que la información cuantitativa es válida siempre y cuando se matice con instrumentos y información más cualitativa. Un docente nos explica que considera esencial para alcanzar una evaluación justa:

Darle más peso a la evaluación continua, pero por otro lado considero también que es necesario hacer un examen, entonces me parece una pregunta difícil, pero sí creo que el tema de la evaluación continua es muy importante porque al final vas viendo el trabajo diario de esas personas, el esfuerzo, la progresión y al final un examen pues bueno, hay gente que funciona muy bien en los exámenes y tiene suerte. [D9,SEC,CPU,DES:1]

Siguiendo esta idea, otro docente concibe que:

No creo en los exámenes, estresan mucho a los alumnos, no mides lo que realmente saben. Lo justo es realizar una evaluación continua por observación día a día. No me convence que se haga cada dos semanas un examen escrito, porque al final nos terminamos apoyando en el examen escrito y ahí fallamos, y los estudiantes aprenden de que lo único que cuenta es el examen y en estudiar memorísticamente para el examen, sin aprender. [D23,PRI,CPU,DES:4]

Asimismo, los docentes destacan que la evaluación sea procesual, es decir, continua y que mida el progreso real del estudiante:

Para llevar cabo una evaluación justa es importante realizar una evaluación continua pero no evaluación continua como lo que se suele hacer de voy poniendo positivos cada día sino voy viendo cómo evoluciona en alumno, si ha mejorado le valoro que ha mejorado. Una persona que estaba en un 1 si ha sacado un 3 para mí eso es una evolución o ha sacado un 5. [D16,SEC,CC,DES:1]

No cabe duda que existe una radical diferencia en lo que se considera una evaluación justa entre los distintos docentes, teniendo profundizas implicaciones que la forma que tienen de entender la evaluación y por ende la educación.

Este grupo de docentes que tiene una concepción equitativa de la evaluación justa se define por encontrarse mayoritariamente en centros situados en contextos socio-económicamente desafiantes, sin encontrar diferencias entre la titularidad o el nivel educativo que imparten.

Como panorámica global de lo encontrado, en la tabla II se resumen las diferentes concepciones de evaluación justa halladas. Asimismo, se aportan datos complementarios para

el análisis de las ideas implícitas del profesorado.

Evaluación justa igualitaria	Evaluación justa equitativa
Explicitar los criterios	Adaptar los criterios según las necesidades de cada estudiante
Objetividad	Subjetividad a favor de cada estudiante
Transparencia y explicitación de los criterios	Individualización de los criterios de evaluación
Rendimiento	Progreso y esfuerzo
Preparación de pruebas estandarizadas	Preparación para la vida
Examen como elemento clave de evaluación	Restar peso al examen en la evaluación
Uso de pruebas cualitativas solo como complemento del examen	Uso de pruebas cualitativas de evaluación
Evaluación de los contenidos	Evaluación de las actitudes y valores
Evaluación igual para todos los estudiantes	Evaluación adaptada a cada uno de los estudiantes

Tabla II. Comparación entre las concepciones docentes sobre una evaluación justa

5. Discusión y conclusiones

En el estudio presentado se han obtenido evidencias empíricas de las percepciones de los docentes acerca de qué es una evaluación justa. Esta línea de investigación, poco estudiada hasta el momento, aporta evidencias sobre las ideas implícitas de los docentes acerca de la evaluación y cómo se traducen éstas en su forma de evaluar a los estudiantes.

Los resultados de las entrevistas realizadas a los docentes muestran que existe un elemento clave en las diferentes concepciones de los docentes sobre qué es una evaluación justa (igualitaria o equitativa): el contexto en el que se encuentra en centro educativo en el que trabajan. Así, los docentes que trabajan en entornos favorables consideran que una evaluación justa es aquella objetiva, cuantitativa y transparente en cuanto proceso. No obstante, para los profesores que enseñan en centros ubicados en contextos desafiantes y que tienen una diversidad de estudiantes mayor, indican que la evaluación justa debe adaptarse a las necesidades de ellos, ser más cualitativa y medir el progreso real de los alumnos.

El estudio de Brown, Hui, Yu y Kennedy (2011) es un referente al respecto. Allí se destaca la importancia del contexto en la evaluación de los estudiantes: la cultura así como el status social de los docentes inciden en la forma en que éstos conciben la enseñanza y la evaluación. Estos autores encontraron que mientras que algunas concepciones educativas pueden ser esencialmente iguales en toda la población, las concepciones sobre la evaluación son legítimamente diferentes por las diferencias sociales y culturales existentes en las escuelas y en la sociedad.

Otro elemento encontrado que no incide en la forma de concebir que es una evaluación justa es el nivel educativo en el que ejerzan los docentes. Así, no hemos encontrado diferencias en cuanto a la concepción más igualitaria o equitativa de la evaluación por ser profesor de Educación Primaria o Secundaria.

Si relacionamos estos hallazgos con la propuesta de Lam (1995) vemos que los resultados coinciden con los docentes que tienen una perspectiva más cercana a la equidad. Así, tanto Lam (1995) como los docentes entrevistados consideran que una evaluación es injusta si los

estudiantes no tienen las mismas igualdades de oportunidades para demostrar aquello que conocen, si las evaluaciones son sesgadas y solo sirven para evaluar un tipo de estudiantes o alguna capacidad en específico y por último los docentes toman decisiones en detrimento de aquellos estudiantes que presentan mas dificultades.

Los resultados de nuestra investigación también están en concordancia con Suskie (2000) la cual considera que para lograr una evaluación justa es necesario: a) establecer claramente las metas y los resultados de aprendizaje; b) unir la evaluación con aquello que se enseña y viceversa; c) utilizar distintos instrumentos y pruebas de evaluación así como evaluar en diferentes momentos; d) ayudar a los estudiantes a aprender a realizar satisfactoriamente las pruebas de evaluación; e) empoderar a los estudiantes; y f) interpretar los resultados de la evaluación de forma adecuada.

No cabe duda de que este humilde estudio apenas ha podido asomarse a una temática tan compleja como desafiante como es la evaluación justa, por lo que sería muy interesante seguir investigando sobre el tema. Este trabajo ha abierto una interesante línea de investigación relacionada con la incidencia del contexto socio-económico en las concepciones sobre evaluación de los docentes. Y las conclusiones son claras: dicho contexto determina las concepciones y con ello la forma de enseñar de los docentes, y no está exenta de una forma de ver la educación (Autores, 2015a).

Con esta investigación se pretende contribuir a un debate informado en España acerca del impacto que están produciendo los cambios en la evaluación de los estudiantes. Se han aportado evidencias acerca de que la imposición de pruebas externas estandarizadas con carácter certificador choca con las concepciones de lo que es una evaluación justa por parte de muchos documentos españoles. Esto puede generar un desajuste cuyas consecuencias recaigan directamente en los estudiantes. Parece necesario que se abra un debate sobre qué es una evaluación justa y sus implicaciones.

Las aportaciones novedosas de este artículo son en primer lugar, en el uso de la fenomenografía para profundizar en las concepciones de cómo tiene que ser una evaluación justa, muy poco utilizado en este ámbito especialmente en los estudios españoles. En segundo lugar, la temática de una evaluación justa ha sido poco investigada hasta el momento desde un punto de vista empírico, por lo que nuestro estudio aporta datos útiles para el avance del conocimiento y para mejorar la educación.

Esta investigación, como cualquier otra, no está exenta de limitaciones. Además de las propias de un estudio fenomenográfico con pocos participantes y quizá no representativos de los docentes españoles, hay que destacar dos claros puntos débiles. En primer lugar, destacamos que las concepciones ambiguas, contradictorias o cambiantes de muchos docentes hacen que no haya sido posible ubicar claramente a todos los docentes en un grupo u otro, con lo que, en segundo lugar, esta trabajo se muestra incapaz de determinar qué hace que los docentes tengan una concepción más igualitarista o más equitativa de lo que es una evaluación justa.

Para complementar nuestro estudio, sería interesante realizar un estudio cuantitativo ex post facto con la misma temática, para conocer de una forma más amplia las concepciones sobre evaluación justa de los docentes. Asimismo, se podría llevar a cabo un estudio tanto cuantitativo ex post facto como cualitativo con estudiantes para conocer sus concepciones de

evaluación justa, pudiéndola comparar con las aportadas por los docentes.

La principal conclusión de este estudio es la necesidad de abrir un debate entre docentes, administradores, tomadores de decisiones e investigadores sobre qué es una evaluación justa y, con ello, cómo es la evaluación de los estudiantes. Desde nuestro punto de vista, si se opta por un modelo igualitario, con pruebas “objetivas” y sin diferencia se estará apostando por una educación centrada en la selección y jerarquización de los estudiantes, con lo que se profundiza en las desigualdades sociales. Si se opta por un modelo basado en la equidad, adaptado a las necesidades de los estudiantes, participativo, con instrumentos variados, cualitativo y esencialmente formativo, se estará apostando por una educación que contribuya a una sociedad más justa.

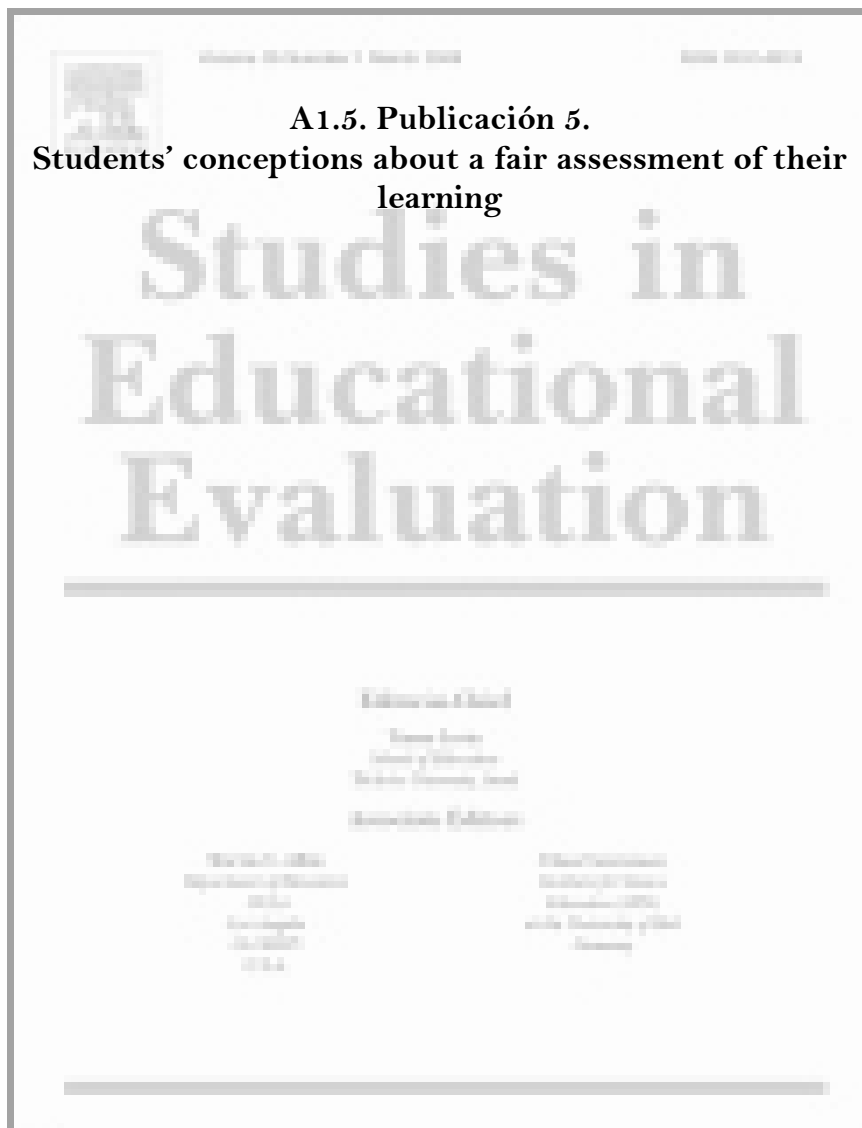
En conclusión, el debate sobre qué es una educación justa así como las implicaciones que tiene en la práctica docente es una temática fundamental a abordar si queremos una evaluación que luche para lograr una sociedad más justa y equitativa (Autores, 2015b).

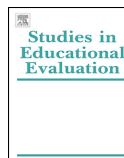
Referencias

- Åkerlind, G. (2002). Principles and Practice in Phenomenographic Research. Paper presented at the *Current Issues in Phenomenography Conference*, Canberra, ACT.
- Andrade, H. y Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32(2), 159-181. doi: [10.1080/02602930600801928](https://doi.org/10.1080/02602930600801928).
- Autores (2015a).
- Autores (2015b).
- Autores (2014).
- Azis, A. (2015). Conceptions and practices of assessment: a case of teachers representing improvement conception. *Teflin Journal*, 26(2), 129-154. doi: [10.15639/teflinjournal.v26i2/129-154](https://doi.org/10.15639/teflinjournal.v26i2/129-154)
- Bowden, J. (1996). Phenomenographic research - Some methodological issues. En G. Dall'Alba y B. Hasselgren (Eds.), *Reflections on Phenomenography: towards a methodology* (pp. 49-66). Suecia: Acta Universitatis Gothoburgensis.
- Bowden, J.A. y Walsh, E. (2000). Phenomenography. *Phenomenography*, 1, 154-169.
- Brown, G. T. L. (2003). *Teachers' conceptions of assessment*. Tesis doctoral, Universidad de Auckland.
- Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged version 1, 2. *Psychological Reports*, 99(1), 166-170. doi: [10.2466/pr0.99.1.166-170](https://doi.org/10.2466/pr0.99.1.166-170).
- Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. Auckland: Nova Science Publishers.
- Brown, G. T. L., Hui, S. K., Flora, W. M. y Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5), 307-320. doi: [10.1016/j.ijer.2011.10.003](https://doi.org/10.1016/j.ijer.2011.10.003)
- Brown, G. T. L. y Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education*, 2, 1-19.

- Brown, G. T. y Remesal, A. (2012). Prospective teachers' conceptions of assessment: A cross-cultural comparison. *The Spanish journal of psychology*, 15(01), 75-89. Doi: [10.5209/rev_SJOP.2012.v15.n1.37286](https://doi.org/10.5209/rev_SJOP.2012.v15.n1.37286)
- Cole, N. S. y Zieky, M. J. (2001). The new faces of fairness. *Journal of Educational Measurement*, 38(4), 369-382. doi: [10.1111/j.1745-3984.2001.tb01132.x](https://doi.org/10.1111/j.1745-3984.2001.tb01132.x).
- Coll, C. y Remesal, A. (2009). Mathematics teachers' conceptions about the functions of assessment in compulsory education. *Journal for the Study of Education and Development*, 32(3), 391-404. doi: [10.1174/021037009788964187](https://doi.org/10.1174/021037009788964187).
- Gebril, A., & Brown, G. T. (2014). The effect of high-stakes examination systems on teacher beliefs: Egyptian Teachers' Conceptions of Assessment. *Assessment in Education: Principles, Policy & Practice*, 21(1), 16-33.
- Gipps, C. y Stobart, G. (2009). Fairness in assessment. En C. Wyatt-Smith y J. Cumming (Eds.), *Educational assessment in the 21st century* (pp. 105-118). Amsterdam: Springer.
- Griffiths, T., Gore, J. y Ladwig, J. (2007). Teachers' fundamental beliefs, commitment to reform and the quality of pedagogy. In Australian Association for Research in Education (Ed.), *Proceedings of the AARE 2006 Conference* (pp. 1-38). Adelaide, Australia: Universidad de Newcastle.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning—tensions and synergies. *Curriculum Journal*, 16(2), 207-223. doi: [10.1080/09585170500136093](https://doi.org/10.1080/09585170500136093).
- Harris, L. R. (2008). A phenomenographic investigation of teacher conceptions of student engagement in learning. *The Australian Educational Researcher*, 35(1), 57-79. doi: [10.1007/BF03216875](https://doi.org/10.1007/BF03216875)
- Harris, L. y Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practices*, 16(3), 365-381. doi: [10.1080/09695940903319745](https://doi.org/10.1080/09695940903319745).
- Lam, T. (1995). *Fairness in Performance Assessment*. Recuperado de <http://www.ericdigests.org/1996-4/fairness.htm>
- Marton, F. (1986). Phenomenography: A research approach to investigating different understandings of reality. *Journal of Thought*, 21(3), 28-49.
- Marton, F. y Booth, S.A. (1997). *Learning and awareness*. Hillsdale, NJ: Lawrence Erlbaum Ass.
- Pepper, M. B. y Pathak, S. (2008). Classroom contribution: What do students perceive as fair assessment?. *Journal of Education for Business*, 83(6), 360-368. doi: [10.3200/JOEB.83.6.360-368](https://doi.org/10.3200/JOEB.83.6.360-368).
- Pettifor, J. L. y Saklofske, D. H. (2012). Fair and ethical student assessment practices. En C. F. Webber y J. L. Lupart (Eds.), *Leading student assessment* (pp. 87-106). Amsterdam: Springer.
- Picón Jácome, É. (2013). La rúbrica y la justicia en la evaluación. *Íkala, Revista de lenguaje y cultura*, 18(3), 79-94.
- Prieto, M. y Contreras, G. (2008). Las concepciones que orientan las prácticas evaluativas de los profesores: un problema a develar. *Estudios Pedagógicos (Valdivia)*, 34(2), 245-262.
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, 27(2), 472-482. doi: [10.1016/j.tate.2010.09.017](https://doi.org/10.1016/j.tate.2010.09.017)
- Sambell, K., McDowell, L. y Brown, S. (1997). "But is it fair?": An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, 23(4), 349-371. doi: [10.1016/S0191-491X\(97\)86215-3](https://doi.org/10.1016/S0191-491X(97)86215-3)

- Scott, S., Webberb, C., Lupartc, J., Aitkend, N. y Scott, D. (2014). Fair and equitable assessment practices for all students. *Assessment in Education: Principles, Policy & Practice*, 21(1), 52-70. doi: [10.1080/0969594X.2013.776943](https://doi.org/10.1080/0969594X.2013.776943)
- Sjostrom, B. y Dahlgren, L. O. (2002). Nursing theory and concept development t or analysis: Applying phenomenography in nursing research. *Journal of Advanced Nursing*, 40(3), 339-345.
- Stobart, G. (2005). Fairness in multicultural assessment systems. *Assessment in Education*, 12(3), 275-287. doi: [10.1080/0969594050033724](https://doi.org/10.1080/0969594050033724)
- Suskie, L. (2002). Fair assessment practices: Giving students equitable opportunities to demonstrate learning. *Adventures in Assessment*, 14, 5–10.
- Tierney, R. D., Simon, M. y Charland, J. (2011). Being Fair: Teachers' Interpretations of Principles for Standards-Based Grading. *The Educational Forum*, 75(3), 210-227. doi: [10.1080/00131725.2011.577669](https://doi.org/10.1080/00131725.2011.577669)
- Trigwell, K. (2000). A phenomenographic interview on phenomenography. *Phenomenography*, 1, 62-82.
- Van den Berg, B. (2002). Teachers' meanings regarding educational practice. *Review of Educational Research*, 72, 577-625. doi: [10.3102/00346543072004577](https://doi.org/10.3102/00346543072004577)





Students' conceptions about a fair assessment of their learning



F. Javier Murillo^a, Nina Hidalgo^{b,*}

^a Autonomous University of Madrid, Faculty of Teacher Training and Education, Department of Didactics and Theory of Education, Office III-302, Tomas y Valiente Street, 3 Colmenar Road Km. 15, 28049 Madrid, Spain

^b Autonomous University of Madrid, Faculty of Teacher Training and Education, Department of Didactics and Theory of Education, Office II-102.2, Tomas y Valiente Street, 3 Colmenar Road Km. 15, 28049 Madrid, Spain

ARTICLE INFO

Article history:

Received 23 September 2016

Received in revised form 13 December 2016

Accepted 15 January 2017

Available online xxx

Keywords:

Assessment

Fairness

Pupils' conceptions of assessment

Phenomenographic research

ABSTRACT

The objective of this study is to know the conceptions of primary and secondary school students about what constitutes a fair assessment. To achieve this aim, a phenomenographic study was carried out with 32 students. In a broad sense, the results showed that there were two types of conceptions. On one hand, a conception of fair assessment related to equality, transparency, objectivity, and evaluation of class content; that is, an egalitarian conception of fair assessment. On the other hand, a conception associated with ideas such as adaptation, diversification of tests, and qualitative assessment, even taking into account students' effort and attitudes, which define a conception of fair assessment linked to equity. These conceptions are associated with the concepts of legal justice and social justice, respectively and they contrast with the nature of external evaluations currently carried out.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

It is not very risky to affirm that the assessment of students is one of the activities with worst reputation within the educational system. This is what experience tells us and what is reflected in research, such as in the classic study by Sambell, McDowell, and Brown (1997). In this paper the authors found that teachers and students had very negative perceptions about the practice of assessment, both because of the way in which the assessment process unfolded and because of its implications; they essentially related the practice of assessment with a search for accountability, in which assessment functioned to certify, classify, and label students.

Perhaps for this reason, or perhaps because of the importance assessment has in the process of teaching and learning and the influence conceptions have on its development, in recent years a line of empirical research seeking to advance the understanding about teachers' and students' conceptions of assessment has gained strength. The reason is clear: if we expand our knowledge about the conceptions students have about assessment, we will be in better condition to develop better assessments, those that most closely match their needs and expectations, and thus help improve their learning (Murillo & Hidalgo, 2015a; Brown, 2008; Ibarra-Sáiz

& Rodríguez-Gómez, 2014; Rodríguez-Conde, Olmos-Migueláñez, & Nieto-Isidro, 2014).

In this study we focused on examining the conceptions of students in regards to a key element: what constitutes, for them, a fair assessment (Tierney, 2013). It seems reasonable to think that if students consider that assessments are unfair, the influence these may have on their subsequent learning may be considerably lower than if they think otherwise. With this in mind, we designed a phenomenographic study with students from primary and secondary education, with different characteristics and from different contexts.

2. Theoretical framework

We define conceptions as a set of beliefs originated in the experiences of each person and entrenched in the social environment through interactions (Feixas, 2010; Hasselgren & Beach, 1997; Marton, 1981; Säljö, 1994). In great part they are constructed through the interaction with other people, thus the way in which we conceive the world is directly influenced by the individuals around us.

Within the broad area of research on students' conceptions within the educational field, many researchers have specifically focused on their conceptions about assessment (Tierney, Simon, & Charland, 2011). Although, among all, the work carried out by Professor Galvin Brown and his team at the University of Auckland for more than a decade deserves recognition (Brown & Harris,

* Corresponding author.

E-mail addresses: javier.murillo@uam.es (F. J. Murillo), nina.hidalgo@uam.es (N. Hidalgo).

2012; Brown & Hirschfeld, 2008; Brown, 2008; Brown, 2011; Brown, Irving, Peterson, & Hirschfeld, 2009; Brown, Kennedy, Fok, Chan, & Yu, 2009; Harris, Brown, & Harnett, 2014; Hirschfeld & Brown, 2009; Peterson & Irving, 2008).

In research by Brown and Hirschfeld (2008), for example, an ex post facto study was conducted with 3469 high school students in order to know the students' conceptions about assessment. Factor analysis showed that students had four different types of conceptions: (a) assessment as a way to improve performance and learning, (b) assessment makes students accountable and responsible, (c) assessment is irrelevant, and (d) assessment is enjoyable. The authors concluded that the first conception positively correlated with performance, while students' conceptions b, c and d have a negative correlation.

In the same vein, Peterson and Irving (2008) inquired about the conceptions that high school students had about assessment and its communication. The study, of qualitative nature, involved five focus groups in four schools in New Zealand. The results indicated that students conceived assessment as useful when it served to enhance their learning, and irrelevant when it did not provide them with useful information. The authors concluded that their study offered support from a qualitative standpoint to the findings of the previous quantitative studies by Brown et al.

In a later study, Brown, Irving et al. (2009) added that students' conceptions were closer to assessment as a training practice and not as practice of teacher control. Shortly afterward, Brown and Harris (2012) complemented these results by carrying out a study aiming to know whether there were differences in the conceptions of students based on their level of education. The results showed that students in secondary education, in relation to those in primary education, scored higher in conceiving assessment as irrelevant in comparison to conceiving it as external, as a tool for improvement, or as pleasant. However, in general, all students, without significant differences, conceived assessment mostly as a means to help improve their learning.

Segers and Tillemans (2011), continuing with this line of research, conducted a study with 712 students in the Netherlands with the objective to know their conceptions about the purpose of assessment. Factor analysis suggested that there were five conceptions about the purpose of assessment: (a) it is beneficial for learning, (b) it serves as an instrument to measure what students have learned, (c) it is pleasant to students, (d) it improves the climate of support and collaboration within the classroom, and (e) it serves to exert accountability. The authors noted that students who perceived assessment as an instrument to exert accountability associated it to a profoundly unfair practice.

This issue of the meaning of a fair or unfair assessment has been acquiring special importance in recent years. However, even though the theoretical contributions on the subject multiply (eg, Cole & Zieky, 2001; Gipps & Stobart, 2009; Pepper & Pathak, 2008; Pettifor & Saklofske, 2012), few empirical studies have so far been carried out to explore this subject (Tierney, 2013).

One of the first studies examining students' conceptions about fair assessment was carried out by Sambell et al. (1997) in the United Kingdom. Their objective was to learn about the perceptions of fair assessment from the perspective of students and determine the impact of the validity of such evaluative practice. The methodology used was the development of thirteen case studies to examine alternative methods of assessment and the validity thereof through interviews with students and teachers, observation, and review of relevant assessment materials. The main findings of the study showed that: (a) traditional assessment methods interfered with the implementation of alternative assessment methods, which were perceived as a fairer process for student learning, (b) final assessments, especially through examinations, were deeply unfair and did not reflect everything

students had acquired and learned (c) students claimed that for an evaluation to be fair, it had to be close to their lives and help them develop skills used in their daily lives. Knowledge considered little useful by students was perceived as unfair. As a final conclusion, the authors pointed out that students did not refer to the term "validity" when talking about a fair assessment, but rather to impartiality, a concept closer to equity in assessment.

More than a decade later, Pepper and Pathak (2008), from the Southwestern University (Texas), conducted a study that sought to examine students' perceptions of justice about how the practice of fair assessment contributes to the methods used in assessment in higher education. To this end, a double quantitative study was carried out. The first of these studies consisted in the application of a scale adapted from the *Selection Procedural Justice Scale* to determine the attitude of students towards Evaluative Justice. The second study was experimental in nature, and included two groups of college students. Through simulation exercises, the six conditions that may influence fair assessment were implemented in both groups: assessment explicitness (high or low), pro-activity (high or low) and feedback (high or low). The results of both studies showed that explicitness in assessment and grading criteria, frequent feedback, and proactivity in the assessment process, led students to conceive assessment as fair. Also, from the perspective of students, the explicitness of the criteria and feedback were the two most influential aspects in the development of a fair assessment practice.

One of the most interesting studies in this area, among those conducted to date is that by Scott, Webber, Lupart, Aitken, and Scott (2014). These authors followed a mixed design approach (quantitative and qualitative) in order to examine the conceptions of what is a fair assessment. Participants were 3312 teachers, students, family members, and senior management staff. The results showed that fair assessment practices were closer to the concept of equity and offered five key ideas about what a fair assessment should be like: (a) educators should strive to understand the impact their assessment practices have on students and their families; (b) assessment should be individualized, to meet the capacity, socioeconomic status, culture, and language of each student; (c) it is necessary that all members of the educational community denounce unsubstantiated and illogical assessment practices; (d) the frequency, intensity, and intrusion of assessments should not overwhelm students or their families; and finally (e) assessment should not be used to counter inappropriate student behaviour or to reward certain desired behaviours.

At a time when external standardized tests are increasingly used worldwide, and consequently both teachers and students are receiving increased pressure, interest in designing systems of assessment that are valid and reliable but also useful and fair has grown. The few studies that have been carried out in this area highlight the importance of the subject, but much remains to know.

Therefore, this research seeks to understand the conceptions that students from primary and secondary education have about fair assessment.

3. Methodology

The methodological approach used in this study was phenomenography. This methodology aims to examine the views of people about a particular phenomenon (Bowden & Walsh, 2000; Marton, 1986; Richardson, 1999), making it especially suited to our research. The result of understanding these concepts is a range of the different ways of thinking and conceptualizing an event. Conceptions are considered as beliefs of human thinking, and if we understand different conceptions we will be able to expand what we know about the psyche of different people.

One of the most important aspects of the phenomenographic approach is the selection of **participants**, in a way that the greatest possible variation of the same conception is guaranteed. It is therefore necessary to determine the aspects that can condition the phenomenon to study and select participants based on these aspects. The sample of participants taking part in this study comprised 32 students selected through a non-probability quota sampling with three selection criteria:

- **Educational level:** Fifteen students were selected from primary education and 17 were selected from mandatory secondary education.
- **Type of educational centers:** Seventeen students attended public schools and 15 attended private centers (private and state-funded).
- **Socio-economic context of schools:** Fourteen students attending schools located in favourable contexts and 18 attending schools in challenging contexts were selected. The indicators used to select these centers were: the socio-cultural level of the neighborhood where the center was located, the socio-economic level of the students' families, and the social, cultural, and linguistic characteristics of these students.

The selection of this sample of students was based on accessibility reasons, given that all students attended schools located in the Community of Madrid. The selection procedure was the following; first the most appropriate centers in terms of the three criteria presented above were selected, in a way that we had schools of different types of ownership, educational level, and located in different contexts. Once the schools were selected, we contacted the management staff and sent them a letter with a detailed description of the research. Once the centers knew the goals and needs of our study, the management staff helped us pre-select the students.

Students from the last years of Elementary Education (ISCED 1), 11–12 year-olds, and Compulsory Secondary Education (ISCED 2), 15–16 year-olds, were selected. We consider they reflect the conceptions generated during the educational level that they finish.

Once this first phase of selection was finished, we provided the management staff with a letter to be sent to the students' families, which offered a detailed explanation of the study and asked for signed informed consent for their children to participate, guaranteeing the anonymity and confidentiality of the data. Once the management staff received the consent agreements, between three and four students were chosen from each center, leaving the final distribution of students participating in the study as shown in Table 1.

The **technique used to gather information** was the phenomenographic interview based on self-report (Trigwell, 2000). The interviews started from a generating question: *what is for you a fair assessment?*

From the perspective of Marton (1986) phenomenographic interview is an interactive production where "experiences and conceptions are constructed jointly by the interviewer and the

interviewee" (p. 4427). It is therefore essential that the interview has one or a few pre-determined questions and the researcher always keeps in mind the purpose of the study to guide the interview (Bowden, 1996; Dortins, 2002; Marton, 1986). In this case, the researcher examined the conceptions of students about what is a fair assessment and its implications through the interviews. Also, the researcher should not be positioned against a certain type of conceptions, but be neutral and show shared interest with the study participants.

The interviews had a duration of between 15 and 20 min and were conducted in Spanish. Once completed, the 32 interviews were fully transcribed, faithful to the words of students, and each was assigned a label with the following format [S1: Pri, Pub, Des]. This label contains four pieces of information: (a) the number of student, (b) educational level (Pri: primary education, and Sec: secondary education); (c) type of center (Pu: public, Pr: Private and SF: state-funded) and (d) context where the center is located (Fav: favourable, Cha: challenging).

Once transcribed, an expert the field of assessment translated all interviews from Spanish to English (Temple & Young, 2004). Fieldwork was developed between the months of January and May of 2015.

In accordance with the phenomenographic method, a criterion of proximity to the discourse of the participants was used in accordance with the analysis of the phenomenographic discourse proposed by Marton (1986). The data analysis comprised six stages for the analysis of the conceptions of students. Below are listed each of these six stages of analysis:

- Phase 1. Identification of the words and expressions most used by students relating to their conceptions of a fair assessment.
- Phase 2. Conjugation and organization of selected expressions to create the initial categories or pre-categories of the students' discourse. For their analysis, the principles proposed by Sjöström and Dahlgren (2002) were used: Frequency, Status, and Pregnancy.
- Phase 3. Arrangement of the pre-categories in families so that conceptions remained organized in a tree of meaning consistent with the conceptions of students.
- Phase 4. Selection of the most representative excerpts from each of the pre-categories created in the stage 2, paying special attention to be faithful to the words and meaning of the students' conceptions.
- Phase 5. Comparison of each of the selected excerpts with the other excerpts, creating a pool of meaning of all students' excerpts, allowing us to give a concrete meaning to the excerpts and provide them with a real connotation when comparing them with other relevant fragments about how students conceive a fair assessment (Harris, 2008).
- Phase 6. Determination of the final categories and best excerpts for each of the categories.

4. Results

The analysis of the phenomenographic interviews with students showed that there were two main conceptions about what is a fair assessment: one of them characterized by an egalitarian approach, and the other defined by a more equitable conception. Some students believed that a fair assessment is one that is objective, transparent, and equal for all, while others conceived a fair assessment as one that adapts to the needs and characteristics of different students, and thus, it is more subjective but more useful. A detailed description of these two conceptions is presented next.

Table 1

Distribution of participants according to context, educational level, and the type of ownership of the center.

Context	Primary Education		Secondary Education		Total
	Public S.	Private S.	Public S.	Private S.	
Favorable	3	3	4	4	14
Challenging	5	4	5	4	18
Total	8	7	9	8	32

Source: Designed by the authors.

4.1. Egalitarian fair assessment: "For an assessment to fair all students must be assessed equally"

A group of students proposed that an assessment, to be fair, should be equal for all. From this perspective, they believed that equality in terms of examination, conditions, and support was essential.

A first aspect highlighted by these students was the importance of objectivity in tests. Many of them stated that they sometimes felt that teachers based their grading more on the name of the student than on the content of the test. In other words, students criticized the subjectivity with which some teachers corrected the examinations, stating that teachers needed to be objective. Illustrating this idea, a student indicated that the students' attitude should not be taken into account, but only their performance on the assessments:

For an assessment to be fair all students must be evaluated equally, giving importance to what is done, and also avoiding to take a dislike to some students, for example because of something they may have done at the beginning of the term, and then the teacher takes everything the student does well as a bad thing. That is also somewhat unfair and even if it upset you if the student does well in an exam you have to give him a good grade, instead of judging him for something he did one day. [S12: Sec, Pr, Fav]

The students who conceived assessment from a more egalitarian perspective also believed that the fairest is that the grade depends almost exclusively on the exam, which was seen as the most objective type of assessment to measure their knowledge. Two students indicated:

The assessment must be only exams. More tests, and less attention to behaviour, because that's unfair. [E20: Sec, Pu, Cha]
Exams should be what weights the most, because it is where you put all your knowledge. [S32: Sec, SF, Cha]

Thus, they considered that the exam was the type of assessment that reflected in a faithful and fair way what they had learned. Although they also pointed out that what would be really fair was to have more than one exam; that is, having exams periodically to know in a constant and objective way the process through which students acquire the knowledge established by the teacher. In the words of one student:

Teachers should do three tests per subject, not just a final exam, and take classroom behaviour, notebooks, and exercises more into account ... but giving more weight to the exam, because it is important and shows all your knowledge. [S18: Pri, Pu, Fav]

The exams and conditions must be the same for everyone. Thus, from this egalitarian conception, students considered that assessment must be equal in terms of execution time, type of test, and all conditions related to the application thereof. Otherwise, they insisted, it would be unfair because some would have privileges:

It is only fair that we all have the same test, and if anyone has fewer capabilities then work harder or if not take them to a different classroom and have them do another test equal to all of them. [S8: Sec, Pr, Fav]

It is not fair that different tests are given to students, the tests must be equal for everyone, because we all struggle with different things, I struggle with math and I try hard, and if I do, others can try hard with other things too. [S21: Sec, Pu, Cha]

Finally, students referred to the need for a transparent assessment in regards to two aspects: having prior information and knowing the assessment criteria. Some of the contributions in the phenomenographic interviews indicated that teachers have to offer detailed information about the number of tests and dates,

format, content, duration, the percentage of the final grade each of them represents, and how they will be corrected:

Maybe not in all tests, but they could help us by telling us the questions they will include, or that in some tests we could talk with classmates to write down ideas in common. [S3: Sec, Pr, Fav]

In short, students who had an egalitarian perspective considered that for an assessment to be fair it was necessary that the tests, especially written tests, were objective, transparent, and equal for all.

4.2. Equitable fair assessment: 'A fair assessment must be at the level of students, must be continuous, and use several types of tests or instruments'

The other strong conception that was detected represented a vision of fair assessment based on equity. There were several elements explicit in the phenomenographic interviews associated with this conception.

A first idea highlighted by the students was the need for their participation in the teaching process and, explicitly, in decisions related to their assessment: in regards to the why, how, when ... :

It is essential that we are allowed to participate in class and that they teach us to be more involved in the assessment. [S4: Sec, Pub, Cha]

If we participate, what we have to study sticks easily, we learn more. [S1: Pri, Pub, Cha]

We like to give our opinion in regards to the assessment and we like to choose. For example, it would be fair to give us two types of tests, for example one of multiple choice and another of short answer, because each one has a way of studying and learning, and in assessments we should also be able to show what we have learned in different ways and decide which one we prefer to do. [S9: Sec, Pub, Cha]

A second idea students highlighted was that a fair assessment must measure their efforts and the progress made during the learning process, not just the results. Accordingly, assessments must reflect the real progress shown by students. In the words of three students:

If a student tries hard, has paid attention in class, his attitude is good, has his notebook in good shape, even if he fails but he has made good progress, I would give him a passing grade. [S9: Sec, Pu, Cha]

The exam, in an hour, cannot reflect nine months of course work, the tests are to check, but the teacher has to see the progress of each kid to assign the grades. [S17: Pri, Pub, Fav]

I think that daily work should be taken into account, noticing the effort, because for example, it often seems very unfair that when you have been working hard during the term, and you've been doing all the best you can, and suddenly you maybe have a bad exam and one day you fail. There are some teachers who throw overboard everything we have learned and improved and that's very unfair. [S13: Sec, Pr, Fav]

The students also demanded that the assessment was continuous. They considered that an assessment at the end of the term does not allow students to reflect everything they have learned during that time and their daily class work:

A fair assessment would be a constant assessment, where the day to day has a stronger impact on the grade than a single test, because there is no point that most days you do nothing and that's not taken into account and then you risk it all in a while doing a test and then that is what actually counts for the assessment. [S14: Sec, Pr, Fav]

One more element that students with this conception defended was (a) the availability of several screening tests because a single one cannot capture what has been learned or assess progress, and (b) that these other tests are varied and provide different information.

Indeed, on one hand, the students believed that the exam did not provide enough information to assess the learning acquired in a given subject, a term, or an academic year, as they believed that they knew more than what they could prove in a single examination carried out in an hour:

Not only doing exams, but also other type of tests, because there are classes where you just have to memorize and I really don't think that is useful because after two weeks you forgot everything. So besides an exam, because you have to take into account some data, you have to use other tests to know each other better and how students learn. These tests may be individual assignments where we discuss or reflect on some important issue. [S13: Sec, Pr, Fav]

And, on the other hand, they thought it was very important to use different assessment tools because it helped them demonstrate different skills and abilities that could not be shown in an exam. They thought that the use of alternative tests such as group work, oral tests or debates would help to substantially improve their learning. A student says:

Teamwork has to be evaluated, group assignments, oral examinations, the opinion of the teacher and student, family . . . and that would have to be taken into consideration a little more, because the exams weight more and what we do in class should be what weights the most. [S28: Pri, Pr, Cha]

The adaptation of assessments was seen as another important component to achieve a fair assessment. The students interviewed believed it was unfair that there was one screening test for all even though they were all different:

An assessment is fair if children that struggle the most do a different exam, an assessment at the level of each student, because if a kid for example has problems in math and doesn't understand what you're asking it is fair to give him something easier and we wouldn't get mad. [S1: Pri, Pu, Cha]

The most fair would be to have an individual test for each student and that the teacher takes into account our opinion. [S19: Pri, Pu, Fav]

A group of students believed that a fair assessment from an equity perspective has to focus on their development. Thus, the assessment not only must evaluate the acquisition of cognitive content but also the attitude, coexistence, empathy, and respect for peers. Illustrating this idea:

A fair assessment besides valuing the grades should also assess the ability of being together with each other, because we don't come here just to get high grades, but also to make things better, make life easier for our classmates. [S7: Sec, Pr, Fav]

A fair assessment must be at the level of students and with several tests, because if someone does well on the test but misbehaves with

the classmates or mistreats his friends it is not fair that he has a good grade. Behaviour and companionship have to be assessed, and if we know to work together. [S11: Sec, Pr, Fav]

Another reflection of students about what is a fair assessment was the need for greater teacher involvement in the assessment and that teachers are also evaluated. From the perspective of students, assessment is not about teachers demanding the completion of tasks and then assigning a final grade to students, as this hardly helps or improves learning. The assessment' tasks and tasks required by teachers have to go hand in hand with greater feedback, for the assessment to help students to know how and what to improve. One student reflected this idea as follows:

A fair assessment would have to leave enough time for students to study and review the class contents with the teacher, but above all that the teacher helps students and becomes interested, and does not say, I will leave you with these 40 exercises for you to study, instead the teacher should help us and give us feedback by telling us that we have done well and what we need to improve and how. [S17: Pri, Pu, Fav]

A final element that stands out from the interviews is that students thought that assessments would have to stop being so fragmented into subjects and become more integrated. They pointed out that when leaving school the knowledge they acquired will have to brought together in their day to day. From this perspective, they wondered why teaching and assessing the class contents in such a disconnected way, or why assessing mathematics and social science in isolation if in real life such knowledge will have to be used together. In the words of one of them, for an assessment to be fair it would have to:

See how we progress in general, not in every class separately, because after all what we learn is all related and we are the same student in one subject or another. [S21: Sec, Pu, Cha]

As a global view of the findings, [Table 2](#) summarizes the different conceptions of fair assessment presented by students as a synthesis of the different topics discussed in the Results section.

Interestingly, there does not seem to be any clear trends in the conceptions of students depending on the context where the schools were located, their educational level, or the type of school.

5. Discussion

The two conceptions of fair assessment that were proposed by the students (egalitarian versus equitable) are related to the two meanings of the adjective "fair". On one side it is associated with legal (or general) justice and on the other with distributive (or social) justice. Indeed, some students related assessment with the establishment of a value judgment and as such, it should be based on the principle that "all human beings are equal before the law, without any privileges or prerogatives of any kind", which is known as legal equality. On the other hand, other students have conceptions of fair assessment closer to Social Justice; that is,

Table 2
Comparison between the students' conceptions of fair assessment.

Egalitarian Fair Assessment	Equitable Fair assessment
Making criteria explicit	Adapting the criteria to the needs of each student
Objectivity	Subjectivity in favor of each student
Transparency and explicitness of the criteria	Individualized assessment criteria
Performance	Progress and effort
Examination as a key element of assessment	Reducing the weight of the exam in the assessment
Use of qualitative tests only as a complement of exams	Use of qualitative tests in the assessment
Assessment of class contents	Assessment of attitudes and values
Equal assessment for all students	Assessment adapted to each student

Source: Designed by the authors.

based on the principle of redistribution, sensibility, and differential support (Murillo & Hidalgo, 2015b).

This binomial of equality versus equity found in the students' conceptions of fair assessment is not new in the literature. Thus, Lam (1995) believes that ensuring equality in assessment allows for the comparison of results and the simplification of assessments, but its significance is lost, ceasing to be impartial; while ensuring fairness reduces bias and leads to a meaningful assessment but introduces difficulties in the administration and comparison of the students' results. According to this author, a fair assessment from an egalitarian approach focuses on ensuring equal conditions in the test administration (materials, time, space, resources . . .). An egalitarian fair assessment requires the same administration, content, grading, and interpretation of results, therefore aiming to achieve objectivity.

Opposite to this position, authors like Suskie (2002) argue that a fair assessment is not one that is the same for all students, as not all students have had necessarily the same opportunities of access to education; so the principle of equality is not enough to ensure justice. This author considers that is essential to promote equity in assessment to ensure justice, adapting assessment to the needs and characteristics of students. Such equity translates into carrying out multiple assessments with different instruments to ensure that the diversity of students is properly represented through the assessment.

This study presents some elements in common with a study by Scott et al. (2014). In congruence with their results, from the conception of equity in regards to the fair assessment of students, they suggest that assessment should be tailored to the needs and characteristics of each student and it should be comprehensive. Also, Scott et al. (2014) emphasize the need for objectivity and transparency in the assessment process, also highlighted by students in our study who conceived assessment from an egalitarian perspective.

Specifying assessment criteria and grading as well as frequent feedback are the highlights of students perspective in Pepper and Pathak's (2008) research. Transparency in assessment criteria is also highlighted in our study while constant feedback is not mentioned by our participants as an element that makes assessment practice a fairer process.

It is interesting that no differences in conceptions were found between students from different educational, contexts, and type of schools. This analysis contrasts with a study conducted previously with teachers, where differences in conceptions were found depending on the context where the schools were located (Murillo, Hidalgo, & Flores, 2016). Other studies related to assessment conceptions of both teachers and students point out the influence of context on implicit beliefs about assessment practice (Brown, Hui, Flora, & Kennedy et al., 2011; Davis & Neitzel, 2010; James & Pedder, 2006; Rubie-Davies, Flint, & McDonald, 2012; Vandeyar & Killen, 2007). It seems clear that more studies need to be conducted to understand what factors have an effect on these very different conceptions.

6. Conclusions

It is not very risky to affirm that one of the elements that has characterized education the early years of this century is the overdevelopment of assessment in all areas. In a few years we have gone from requesting the promotion of a culture of evaluation to an obsession to assess everything and at all times. And the ones suffering such pressure the most, as always, are the students, given they are the weakest link in the educational system.

The over-emphasis given to international assessment and the increase in quasi-market educational systems in many countries around the world with their consequent external assessments to

"inform" families, have led in recent years to a growing interest in knowing what is a fair assessment and, above all, the ways to develop them. In that context, fair assessment has been too often associated with terms such as objectivity, reliability, transparency, or validity. Not in vain these are the characteristics required by external assessments, both national and international. That is, we are witnessing a development of the egalitarian conception of fair assessment.

Faced with this reality, we find that students in primary and secondary education have two positions about what is a fair assessment, regardless of the type of school they attend or the context where the school is located. The first one supports a more elaborate conception of fair assessment, closer to Social Justice and to the importance that education contributes to compensate inequalities. The other conceptions are aligned with the pressure to consider a fair assessment as objective and equal for all. How would the first group of students face that discrepancy between their conception and the external reality?

This study raises more questions than it answers. What are the characteristics of students in the two groups? Is this related to their age? Is there an evolution in the concept of fair assessment? Is it influenced by their teachers? Is there any relationship with the conceptions of teachers or parents about this issue? . . . But above all, what are the implications that one conception or another can have on their learning?

The methodological feature of phenomenographic research followed in this study consisting in having a wide variety of participants has not allowed to detect regularities to help identify the characteristics of the two groups. This however was not the objective, because we did not expect to find these two positions so clearly opposed. This calls for more studies that contribute to provide more information. In any case, we can consider this study a good first approximation that shall open the field to future studies.

Two final ideas constitute the legacy of this study and we want to finish by making them explicit. On one hand, the existence of two conceptions of fair assessment in students that are ontologically apposed: an egalitarian conception, linked the concept of legal justice, and another conception of fair assessment based on equity, related to social justice. The origin and consequences of these two conceptions remains to be studied.

The other lesson learned is the very elaborated concept of fair assessment that these very young students have shown. In these times in which competitiveness and quasi-market ideas are permeating our educational systems, our schools, and many of their teachers and scholars, it results encouraging that a good number of students supports an assessment linked to a comprehensive and equitable educational development. Hopefully we will learn from these approaches, hopefully we will not end up with these ideas, but they will help us build fairer schools. Because there is no alternative, only with fair schools can we contribute to fair societies. John Rawls already said it (1971), if a social institution is not fair above all other considerations . . . what's the use? Students have shown to be aware of this; it is time for us to learn from them.

References

- Bowden, J. A., & Walsh, E. (2000). Phenomenography. *Phenomenography*, 1, 154–169.
- Bowden, J. A. (1996). Phenomenographic research: Some methodological issues. In G. Dall'Alba, & B. Hasselgren (Eds.), *Reflections on phenomenography: Towards a methodology* (pp. 49–66). Göteborg: Acta Universitatis Gothoburgensis.
- Brown, G. T. L., & Harris, L. (2012). Student conceptions of assessment by level of schooling: Further evidence for ecological rationality in belief systems. *Australian Journal of Educational & Developmental Psychology*, 12, 46–59.
- Brown, G. T. L., & Hirschfeld, G. H. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3–17. <http://dx.doi.org/10.1080/09695940701876003>.

- Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. Auckland: Nova Science Publishers.
- Brown, G. T. L., Hui, S. K., Flora, W. M., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5), 307–320. <http://dx.doi.org/10.1016/j.ijer.2011.10.003>.
- Brown, G. T. (2011). Self-regulation of assessment beliefs and attitudes: A review of the students' conceptions of assessment inventory. *Educational Psychology*, 31(6), 731–748. <http://dx.doi.org/10.1080/01443410.2011.599836>.
- Brown, G. T. L., Irving, S. E., Peterson, E. R., & Hirschfeld, G. H. (2009). Use of interactive-informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning and Instruction*, 19(2), 97–111. <http://dx.doi.org/10.1016/j.learninstruc.2008.02.003>.
- Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347–363. <http://dx.doi.org/10.1080/09695940903319737>.
- Cole, N. S., & Zieky, M. J. (2001). The new faces of fairness. *Journal of Educational Measurement*, 38(4), 369–382. <http://dx.doi.org/10.1111/j.1745-3984.2001.tb01132.x>.
- Davis, D. S., & Neitzel, C. (2010). The relationship between students' reading orientations and their strategic activity during a collaborative reading task. *Reading Psychology*, 31(6), 546–579. <http://dx.doi.org/10.1080/0270210903256593>.
- Dortins, E. (2002). Reflections on phenomenographic process: Interview, transcription and analysis. *Quality Conversations: Research and Development in Higher Education*, 25, 207–213.
- Feixas, M. (2010). Enfoques y concepciones docentes en la universidad [Teaching approaches and conceptions taken by university faculty]. *RELIEVE: e-Journal of Educational Research, Assessment and Evaluation*, 16(2), 1–27.
- Gipps, C., & Stobart, G. (2009). Fairness in assessment. In C. Wyatt-Smith, & J. Cumming (Eds.), *Educational assessment in the 21st century* (pp. 105–118). Amsterdam: Springer.
- Harris, L. R., Brown, G. T., & Harnett, J. A. (2014). Understanding classroom feedback practices: A study of New Zealand student experiences, perceptions, and emotional responses. *Educational Assessment, Evaluation and Accountability*, 26(2), 107–133.
- Harris, L. R. (2008). A phenomenographic investigation of teacher conceptions of student engagement in learning. *The Australian Educational Researcher*, 35(1), 57–79. <http://dx.doi.org/10.1007/BF03216875>.
- Hasselgren, B., & Beach, D. (1997). Phenomenography: A good for nothing brother of phenomenology? Outline of an analysis. *Higher Education Research and Development*, 16(2), 191–202. <http://dx.doi.org/10.1080/0729436970160206>.
- Hirschfeld, G. H., & Brown, G. T. (2009). Students' conceptions of assessment: Factorial and structural invariance of the SCoA across sex, age, and ethnicity. *European Journal of Psychological Assessment*, 25(1), 30–38.
- Ibarra-Sáiz, M. S., & Rodríguez-Gómez, G. (2014). Modalidades participativas de evaluación: Un análisis de la percepción del profesorado y de los estudiantes universitarios [Participatory assessment methods: An analysis of the perception of university students and teaching staff]. *Revista de Investigación Educativa*, 32(2), 339–362. <http://dx.doi.org/10.6018/rie.32.2.172941>.
- James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109–138. <http://dx.doi.org/10.1080/09585170600792712>.
- Lam, T. (1995). *Fairness in performance assessment*. Greensboro, NC: ERIC Digest.
- Marton, F. (1981). Phenomenography-describing conceptions of the world around us. *Instructional Science*, 10, 177–200. <http://dx.doi.org/10.1007/BF00132516>.
- Marton, F. (1986). Phenomenography: A research approach to investigating different understandings of reality. *Journal of Thought*, 21(3), 28–49.
- Murillo, F. J., & Hidalgo, N. (2015a). Dime cómo evaluas y te diré qué sociedad construyes. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 5–9.
- Murillo, F. J., & Hidalgo, N. (2015b). Enfoques Fundamentales de la Evaluación de Estudiantes para la Justicia Social. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 43–61.
- Murillo, F. J., Hidalgo, N., & Flores, S. (2016). Incidencia del contexto socio-económico en las concepciones docentes sobre evaluación. *Profesorado, Revista de Currículum y Formación del Profesorado*, 20(3), 251–281.
- Pepper, M. B., & Pathak, S. (2008). Classroom contribution: What do students perceive as fair assessment? *Journal of Education for Business*, 83(6), 360–368. <http://dx.doi.org/10.3200/OJEB.83.6.360-368>.
- Peterson, E. R., & Irving, S. E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction*, 18(3), 238–250. <http://dx.doi.org/10.1016/j.learninstruc.2007.05.001>.
- Pettifor, J. L., & Saklofske, D. H. (2012). Fair and ethical student assessment practices. In C. F. Webber, & J. L. Lupart (Eds.), *Leading student assessment* (pp. 87–106). Amsterdam: Springer.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
- Richardson, J. T. (1999). The concepts and methods of phenomenographic research. *Review of Educational Research*, 69(1), 53–82. <http://dx.doi.org/10.3102/00346543069001053>.
- Rodríguez-Conde, M. J., Olmos-Migueláñez, S., & Nieto-Isidro, S. (2014, October). Evaluation in education and guidance. Paper presented in the Second International Conference on Technological Ecosystems for Enhancing Multiculturality. University of Salamanca, Spain.
- Rubie-Davies, C. M., Flint, A. Y., & McDonald, L. G. (2012). Teacher beliefs, teacher characteristics, and school contextual factors: What are the relationships? *British Journal of Educational Psychology*, 82(2), 270–288. <http://dx.doi.org/10.1111/j.2044-8279.2011.02025.x>.
- Säljö, R. (1994). Minding action: Conceiving of the world versus participating in cultural practices. *Nordisk Pedagogik*, 2, 71–80.
- Sambell, K., McDowell, L., & Brown, S. (1997). But is it fair?: An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, 23(4), 349–371. [http://dx.doi.org/10.1016/S0191-491X\(97\)86215-3](http://dx.doi.org/10.1016/S0191-491X(97)86215-3).
- Scott, S., Webber, C., Lupart, J., Aitken, N., & Scott, D. (2014). Fair and equitable assessment practices for all students. *Assessment in Education: Principles, Policy & Practice*, 21(1), 52–70. <http://dx.doi.org/10.1080/0969594X.2013.776943>.
- Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation*, 37(1), 49–54. <http://dx.doi.org/10.1016/j.stueduc.2011.03.008>.
- Sjöström, B., & Dahlgren, L. O. (2002). Applying phenomenography in nursing research. *Journal of Advanced Nursing*, 40(3), 339–345. <http://dx.doi.org/10.1046/j.1365-2648.2002.02375.x>.
- Suskie, L. (2002). Fair assessment practices: Giving students equitable opportunities to demonstrate learning. *Adventures in Assessment*, 14, 5–10.
- Temple, B., & Young, A. (2004). Qualitative research and translation dilemmas. *Qualitative Research*, 4(2), 161–178. <http://dx.doi.org/10.1177/1468794104044430>.
- Tierney, R. D., Simon, M., & Charland, J. (2011). Being fair: Teachers' interpretations of principles for standards-based grading. *The Educational Forum*, 75(3), 210–227. <http://dx.doi.org/10.1080/00131725.2011.577669>.
- Tierney, R. D. (2013). Fairness in classroom assessment. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 125–144). Thousand Oaks, CA: SAGE Publications.
- Trigwell, K. (2000). A phenomenographic interview on phenomenography. *Phenomenography*, 1, 62–82.
- Vandeyar, S. Y., & Killen, R. (2007). Educators' conceptions and practice of classroom assessment in post-apartheid South Africa. *South African Journal of Education*, 27(1), 101–115.

F. Javier Murillo Full professor in the area of Research Methodology and Educational Diagnosis at the Autonomous University of Madrid (UAM), Spain. Coordinator of the Education PhD Program at the UAM. Coordinator of the Educational Change for Social Justice research group (GICE). He was the General Coordinator of the Latin American Laboratory for Assessment of the Quality of Education, UNESCO, and Director of Studies at the Center for Educational and Documentational Research of the Ministry of Education of Spain. He is the coordinator of the Latin American Network for Research on Change and School Effectiveness, Director of the Iberoamerican Journal on Quality, Efficiency and Change in Education, the Iberoamerican Journal of Educational Evaluation, and the International Journal of Education for Social Justice. He has worked as an expert consultant in educational research and assessment in different countries of Latin America, and with various international agencies - UNESCO, OECD and the Convenio Andres Bello.

Nina Hidalgo PhD Candidate in the area of Research Methodology and Educational Diagnosis at the Autonomous University of Madrid, with a doctoral scholarship awarded on a competitive basis by the Ministry of Education of Spain. She holds a Degree in Educational Sciences with a specialization as Early Childhood Education Teacher from the University of Lleida, and a Degree in Psychopedagogy from the Autonomous University of Madrid. She is a member of the Educational Change for Social Justice research group (GICE), and the editorial board of the Journal of Education for Social Justice. She is also editor of the Ibero-American Journal of Educational Evaluation. She has published several articles, recensions, and book chapters.



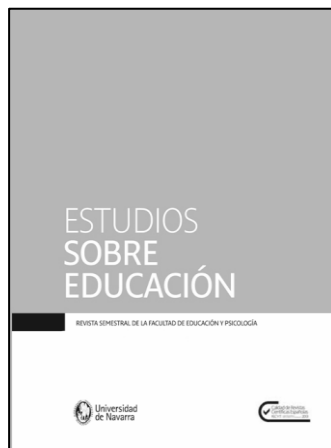
ANEXO 2.

FACTOR DE IMPACTO DE LAS REVISTAS

A continuación se presenta con detalle el factor de impacto de las cinco revistas donde se han publicado cada uno de los cinco artículos que componen esta tesis.

A2.1. Murillo, F. J., Martínez-Garrido, C. e Hidalgo, N. (2014). Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España. *ESE-Estudios sobre Educación*, 27, 91-113. doi: 10.15581/004.27.91-113

El primero de los artículos que compendian esta tesis titulado “*Incidencia de la forma de evaluar los docentes de Educación Primaria en el rendimiento de los estudiantes en España*” fue publicado en junio del año 2014 en la revista Estudios sobre Educación (ISSN vers. impresa 1578-7001, ISSN vers. electrónica 2386-6292). Esta publicación tiene como objeto



primordial contribuir a la difusión del creciente número de resultados de investigación originales, tanto teóricos como experimentales, que se realizan en España y en otros países, especialmente latinoamericanos y del sur de Europa. La publicación es editada por la Universidad de Navarra.

Los artículos publicados en *Estudios sobre Educación* están indexados en importantes bases de datos internacionales de carácter científico como Scopus, Academic Search Premier, Fuente Académica Plus, IBZ Online, International Bibliography of Social Sciences, Public Affairs Index, Education Source, Educational Research Abstracts y en fuentes nacionales como ISOC o Dialnet.

La revista cuenta con el suficiente número de citas para estar en las principales bases de datos que miden el impacto científico:

formó parte del Journal Citation Reports (Web of Science) edición Ciencias Sociales durante los años 2011 y 2012 y desde 2009 a la actualidad está en los índices de impacto SJR (revistas indexadas en Scopus). En 2014, año de publicación del artículo, su SJR era de 0,150 (puesto 775 de 1059 revistas en la categoría Educación). Además, tiene un índice h5 de 9 (posición 41 de 108 revistas en el área de educación) según el Índice H de las revistas científicas españolas según Google Scholar Metrics, 2011-2015.

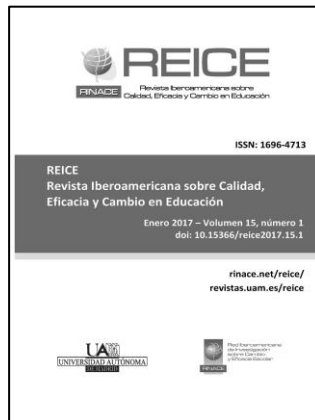
En cuanto a sistemas de clasificación, CARHUS Plus+ ed. 2014 la sitúa en un grupo C mientras que CIRC (ed. 2016) la incluye entre las revistas del grupo B en la categoría de Ciencias Sociales. Latindex la incluye en su catálogo con todos los criterios cumplidos (33): básicos, de presentación, gestión y política editorial y contenido.

En 2010, obtuvo el Sello de Calidad de Revistas Científicas Españolas otorgado por la FECYT (Fundación Española para la Ciencia y la Tecnología). Dicho Sello de Calidad ha sido renovado, con posterioridad, en dos ocasiones: en 2013 y en 2016, con vigencia hasta 2019.

A2.2. Hidalgo, N. y Murillo, F. J. (2017). Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes. *REICE. Revista Iberoamericana de Calidad, Eficacia y Cambio en Educación*, 15(1), 107-128.

doi:10.15366/reice2017.15.1.007

El artículo titulado “*Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes*” (Nina Hidalgo y F. Javier Murillo) ha sido publicado en enero del año 2017 en la REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación (ISSN 1696-4713). REICE publica artículos inéditos que versen sobre los temas de Calidad, Equidad, Eficacia Escolar, Cambio Educativo, Mejora de la Escuela e Innovación Educativa. La revista es editada por la Red RINACE.



La revista REICE tiene un índice h5 de 13 lo que la sitúa en el puesto 18 de 108 revistas en la categoría Educación según el Índice H de las revistas científicas españolas según Google Scholar Metrics, 2011-2015.

En el ranking realizado por la Revista Comunicar se encuentra entre las 25 mejores Revistas de Educación de España. Asimismo en el sistema de clasificación CIRC (Clasificación Integrada de Revistas Científicas), edición 2016, se encuentra en el nivel C en las categorías de Ciencias Sociales y Ciencias Humanas. Latindex la incluye en su catálogo con 34 de los 36 criterios cumplidos para revistas electrónicas.

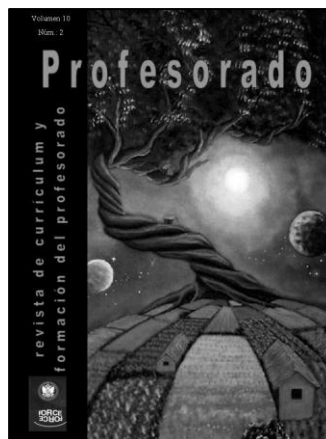
Durante el año 2015 la revista ha pasado a formar parte de la base de datos ESCI (Emerging Sources Citation Index de la Web of Science) y sus artículos también se encuentran indexados en las bases de ISOC (Base de datos del Consejo Superior de Investigaciones Científicas de España) y la Red de Revistas Científicas de América Latina y el Caribe, España y Portugal además de Dialnet o Fuente Académica Plus.

Los artículos publicados en REICE tienen una amplia difusión al encontrarse en índices y catálogos como la Biblioteca de la OEI, IRESIE (Índice de Revistas de Educación Superior e Investigación Educativa), REDIB (Red iberoamericana de Innovación y Conocimiento Científico) y la Red REDIAL, Genamics JournalSeek, RESH (Revistas Españolas de Ciencias Sociales y Humanidades), REBIUN (Red de Bibliotecas Universitarias), Catalogue du Système Universitaire de Documentation y WorldCat y The Elektronische Zeitschriftenbibliothek (Electronic Journals Library).

La revista está comprometida con las políticas de acceso abierto a la ciencia por lo que se encuentra incluida en DOAJ (Directory of Open Access Journals) y Dulcinea.

A2.3. Murillo, F. J., Hidalgo, N. y Flores, S. (2016). Incidencia del contexto socio-económico en las concepciones docentes sobre evaluación. *Profesorado, Revista de Currículum y Formación del Profesorado*, 20(3), 251-281.

La revista *Profesorado. Revista de Currículum y Formación del Profesorado*, editada por la Universidad de Granada (ISSN impreso 1138-414X, ISSN electrónico 1989-639X), es una publicación interdisciplinar de carácter científico-académico, dirigida a investigadores y profesionales de la educación. Pretende fomentar el debate



científico-profesional, el intercambio de ideas y la difusión de resultados de investigación. En esta revista se publicó en diciembre de 2016 el artículo “*Incidencia del contexto socio-económico en las concepciones docentes sobre evaluación*” (F. Javier Murillo, Nina Hidalgo y Sofía Flores).

La revista está indexada en la base de datos de citas ESCI (Emerging Sources Citation Index de la Web of Science) desde 2015 y en Scopus desde 2012. En 2015, su SJR de 0,104 la sitúa en la posición 959 de 1066 revistas de la categoría Educación. Tiene un índice h₅ de 18 lo que la sitúa en el puesto 5 de 108 en el área de educación (Índice h de las revistas científicas españolas según Google Scholar Metrics 2011-2015).

En cuanto a clasificaciones, CIRC (ed. 2016) la sitúa en su grupo B de revistas de Ciencias Sociales y para CARHUS

Plus+ ed. 2014 pertenece al grupo C. En el catálogo Latindex aparece con 34 de 36 criterios cumplidos.

Desde 2013 tiene el sello de calidad FECYT, conseguido en su cuarta edición y con una vigencia hasta 2017.

Respecto a su difusión, la revista Profesorado también se encuentra en las siguientes bases de datos, catálogos y repositorios: a) DOAJ: Directory of Open Access Journal, b) ISOC: Revistas de Ciencias Sociales y Humanidades, c) ULRICH'S: Global Serial Directory, d) IRESIE: Índice de Revistas de Educación Superior e Investigación Educativa, e) AERA: American Educational Research Association, f) SUDOC: Le Catalogue du Système Universitaire de Documentación, g) COPAC: National, Academic and Specialist Library Catalogue, h) ZDB: Zeitschriftendatenbank, i) REBIUN: Red de Bibliotecas Universitarias, j) DIGIBUG: Repositorio Institucional de la Universidad de Granada, k) RedDOLAC: Red Docente de America Latina y el Caribe; REDALYC: Red Innovemos. Innovaciones Educativas por América Latina y el Caribe, y m) Dialnet.

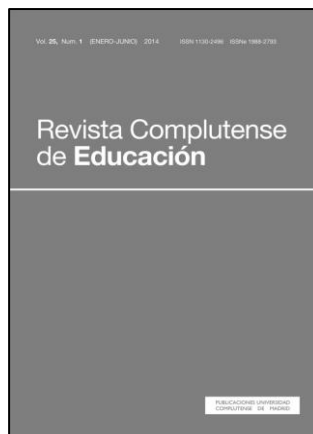
A.2.4. Murillo, F. J. e Hidalgo, N. (en prensa). Las concepciones sobre una evaluación justa de los estudiantes. Un estudio fenomenográfico desde la perspectiva de los docentes. *Revista Complutense de Educación*, en prensa.

El cuarto de los artículos que se presenta compilado en esta tesis doctoral se titula “*Las concepciones sobre una evaluación justa de los estudiantes. Un estudio*

fenomenográfico desde la perspectiva de los docentes” (F. Javier Murillo y Nina Hidalgo) y ha sido aceptado para su publicación en la Revista Complutense de Educación (ISSN impreso 1130-2496, ISSN

electrónico 1988-2793). Esta revista publica artículos relacionados con temas pedagógicos de actualidad desde el punto de vista multidisciplinar e incluye trabajos de investigación y desarrollo tecnológico, de creación e innovación y de descripción y análisis del estado del arte de una cuestión educativa que contribuyan a enriquecer el conocimiento en educación. La publicación es editada por la Universidad Complutense de Madrid.

La *Revista Complutense de Educación* se encuentra indexada en importantes bases de datos internacionales como Emerging Sources Citation Index (Web of Science), Scopus, IBZ Online, Linguistics and Language Behavior Abstracts, PAIS International y en recursos nacionales nacionales como Psicodoc, ISOC o Dialnet.



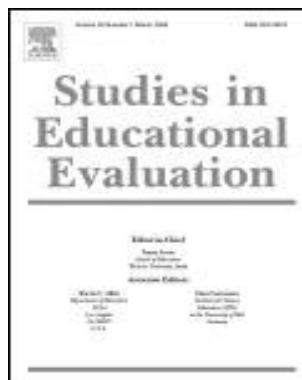
En cuanto a citas en el área de educación, su impacto SJR en 2015 era de 0,167 (posición 739 de 1066 revistas) y su índice h⁵ de 11 (posición 25 de 108 en el Índice h de las revistas científicas españolas según Google Scholar Metrics 2011-2015).

La edición 2016 de CIRC la sitúa entre las revistas del grupo B de Ciencias Sociales, CARHUS Plus+ ed. 2014 en el grupo C y en el Catálogo Latindex aparece con 32 de los 33 criterios cumplidos para revistas impresas.

Desde 2014 cuenta con el sello de la FECYT, conseguido en su 4^a edición y con vigencia hasta el año 2017.

A2.5. Murillo, F. J. e Hidalgo, N. (2017). Students' conceptions about a fair assessment of their learning. *Studies in Educational Evaluation*, 53, 10-17. doi: 10.1016/j.stueduc.2017.01.001

El artículo “*Students' conceptions about a fair assessment of their learning*” (F. Javier Murillo y Nina Hidalgo) será publicado en la revista *Studies in Educational Evaluation* (ISSN impreso 0191-491X, ISSN electrónico 1879-2529) en 2017 en el número 53 de la revista, que saldrá publicado en el mes de junio. Se trata



de una revista que publica informes originales de estudios tanto teóricos como empíricos relacionados con el ámbito de la evaluación educativa. La revista forma parte de la editorial Elsevier y su editor es de la Universidad de Antwerp (Bélgica).

Studies in Educational Evaluation es una prestigiosa publicación indexada en Social Sciences Citation Index (Web of Science) y desde 2015 también en Journal Citation Reports, en cuyo ranking aparece en la posición 175 de 231 revistas de la categoría Educación y en la posición 51 de 57 revistas de la categoría Psicología de la Educación). También está incluida en Scopus desde los años 90 y su impacto SJR en 2015 fue de 0.522 lo que la sitúa en la posición 313 de 1066 revistas de Educación (cuartil 2). Su índice h5 en Google Scholar Metrics es actualmente de 18, es decir al menos 18 artículos publicados con 18 o más citas cada uno.

Asimismo, la revista se encuentra incluida en las siguientes bases de datos, directorios e indexaciones: Academic Search

Premier, ASSIA, Education Source, British Education Index, Contents Pages in Education, Education Research Index, ERIC, Education Abstracts, Education Index, Education Technology Abstracts, Research into Higher Education Abstracts, UnCover, Educational Management Abstracts, Multicultural Education Abstracts, Sociology of Education Abstracts, Technical Education & Training Abstracts y ERA (Educational Research Abstracts Online).



ANEXO 3.

OTRAS PUBLICACIONES Y COMUNICACIONES A CONGRESOS

Paralelamente a las publicaciones que configuran la presente tesis doctoral, se trabajó en la línea de proponer una evaluación alternativa orientada a la Justicia Social. Acerca de esta temática, se intentó contar con publicaciones que llegasen al mayor número de investigadores y docentes posibles, eligiendo revistas de referencia en investigaciones en nuestro contexto.

Entre estas publicaciones contamos, en primer lugar, con *artículos publicados* en distintas revistas:

Hidalgo, N. y Murillo, F. J. (2016). Evaluación de estudiantes para la Justicia Social. Propuesta de un modelo. *Revista Internacional de Educación para la Justicia Social*, 5(2), 159-179. doi:10.15366/riejs2016.5.2.008

Murillo, F.J. e Hidalgo, N. (2015). Enfoques Fundamentantes de la Evaluación de Estudiantes para la Justicia Social. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 43-61.

Murillo, F. J., Hernández-Castilla, R., Hidalgo, N. y Martínez-Garrido, C. (2014). Elaboración y evaluación psicométrica de la escala de actitudes hacia la justicia social en educación (EAJSE). *Revista Internacional de Educación para la Justicia Social*, 3(2), 215-233.

En segundo lugar, también se colaboró en la redacción de un capítulo de un libro de didáctica general

Hidalgo, N. y Murillo, F.J. (2015). Enfoques alternativos de evaluación de los estudiantes. En J. Domingo y M. Pérez Ferra (Coords.), *Aprendiendo a enseñar. Manual Práctico de Didáctica* (pp. 171-174). Madrid: Pirámide.

Por último, se han escrito distintas *editoriales* sobre la temática de una evaluación para la Justicia Social:

Murillo, F.J. e Hidalgo, N. (2016). Hacia un proceso de Evaluación Docente Justo y Socialmente Justo. *Revista Iberoamericana de Evaluación Educativa*, 9(2), 5-8.

Murillo, F.J. e Hidalgo, N. (2016). Evaluación Democrática y para la Democracia. *Revista Iberoamericana de Evaluación Educativa*, 9(1), 5-7.

Murillo, F.J. e Hidalgo, N. (2015). La Evaluación de Estudiantes como Práctica Política y Social en Tiempos

Neoliberales. *Revista Iberoamericana de Evaluación Educativa*, 8(2), 5-7.

Murillo, F.J. e Hidalgo, N. (2015). Dime Cómo Evalúas y Te Diré Qué Sociedad Construyes. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 5-9.

Murillo, F.J. e Hidalgo, N. (2014). Evaluación justa: entre la igualdad y la equidad. *Revista Iberoamericana de Evaluación Educativa*, 7(3), 5-6.

Comunicaciones a congresos

Paralelamente a la publicación de trabajos sobre evaluación para la Justicia Social, la doctoranda también elaboró y presentó comunicaciones junto con otros miembros del grupo de investigación en congresos tanto nacionales como internacionales. Las comunicaciones se agrupan en función de las distintas líneas de investigación trabajadas:

Sobre Concepciones sobre evaluación

Hidalgo, N., Murillo, F.J., Martínez-Garrido, C., Perines, H. y Hernández-Castilla, R. (septiembre, 2015). Teachers' conceptions about a fair student assessment. Comunicación presentada en el congreso *ECER 2015: Education and Transition. Contributions from Educational Research*. Corvinus Budapest, Hungría.

Hidalgo, N. y Hernández-Castilla (junio, 2015). ¿Qué es una evaluación justa? Percepciones de evaluación de los docentes del aprendizaje de los estudiantes. Comunicación presentada en el congreso *AIDIPE 2015*,

*XVII Congreso Internacional de Investigación Educativa:
“Investigar con y para la sociedad, Cádiz, España.*

Referidas a *Evaluación para la Justicia Social*

Hidalgo, N. y Perines, H. (junio, 2016). Evaluación de estudiantes para la Justicia Social. Comunicación presentada en el *XVI Congreso Nacional y VII Congreso Iberoamericano de Pedagogía*, Madrid, España.

Martínez-Garrido, C. e Hidalgo, N. (septiembre, 2013). ¿Cómo evalúan los docentes españoles? Estudio multinivel sobre el impacto de las estrategias de evaluación en el aula sobre el rendimiento académico? Comunicación presentada en *XVI Congreso Nacional y II Internacional de Modelos de Investigación Educativa: Investigación e Innovación Educativa al Servicio de Instituciones y Comunidades Globales, Plurales y Diversas*, Alicante, España.

Sobre *Educación para la Justicia Social*

Hidalgo, N. y Hernández-Castilla, R. (junio, 2016). ¿Cómo es una enseñanza socialmente justa? Un estudio fenomenográfico desde la perspectiva de los estudiantes. Comunicación presentada en el *XVI Congreso Nacional y VII Congreso Iberoamericano de Pedagogía*, Madrid, España.

Perines, H. e Hidalgo, N. (junio, 2016). La cultura de las escuelas para la Justicia Social desde la voz de los estudiantes. Comunicación presentada en el *XVI*

Congreso Nacional y VII Congreso Iberoamericano de Pedagogía, Madrid, España.

Murillo, F, Hernández-Castilla, R., Hidalgo, N., Martínez-Garrido, C. y Perines, H. (septiembre, 2015). Is it possible to measure attitudes toward Social Justice in Education? Development and validation of a scale. Comunicación presentada en el congreso ECER 2015. *Education and Transition. Contributions from Educational Research*, Budapest, Hungría.

Hernández-Castilla y Hidalgo, N. (junio, 2015). Medición de las actitudes de los docentes hacia la justicia social (EAJSE). Instrumento y validación. Comunicación presentada en el congreso *AIDIPE 2015, XVII Congreso Internacional de Investigación Educativa: "Investigar con y para la sociedad*, Cádiz, España.

Murillo, F. J., Hernández-Castilla, R., Hidalgo, N., Paez-Lacadena, P., Cerrillo, R. y Tello, I. (septiembre, 2014). School Culture, key for Social Justice Schools. Comunicación presentada en el congreso *ECER 2014. The Past, Present and Future of Educational Research in Europe*, Porto, Portugal.

Hidalgo, N. y Martínez-Garrido, C. (septiembre, 2013). Estrategias en el aula de atención a la diversidad desde la perspectiva de la Justicia Social. Comunicación presentada en *XVI Congreso Nacional y II Internacional de Modelos de Investigación Educativa: Investigación e Innovación Educativa al Servicio de Instituciones y*

Comunidades Globales, Plurales y Diversas, Alicante,
España.